Department of Signal Processing and Acoustics

OFDM under Oscillator Phase Noise

Contributions to Analysis and Estimation Methods

Pramod Jacob Mathecken

Aalto University

DOCTORAL DISSERTATIONS

Aalto University publication series
DOCTORAL DISSERTATIONS 105/2016

OFDM under Oscillator Phase Noise

Contributions to Analysis and Estimation Methods

Pramod Jacob Mathecken

A doctoral dissertation completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Electrical Engineering, at a public examination held at the lecture hall S2 of the school on 28 June 2016 at 12 o'clock noon.

Aalto University
School of Electrical Engineering
Department of Signal Processing and Acoustics

Supervising professor
Prof. Risto Wichman

Thesis advisor
Dr. Taneli Riihonen

Preliminary examiners
Assistant Prof. Hani Mehrpouyan, Boise State University, USA
Dr. Prathapasinghe Dharmawansa, University of Moratuwa, Sri Lanka

Opponents
Prof. Bhaskar D. Rao, University of California San Diego, USA
Assistant Prof. Hani Mehrpouyan, Boise State University, USA

Aalto University publication series
DOCTORAL DISSERTATIONS 105/2016

ISBN 978-952-60-6838-1 (printed)
ISBN 978-952-60-6839-8 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:978-952-60-6839-8

Unigrafia Oy
Helsinki 2016

Finland

Author
Pramod Jacob Mathecken

Name of the doctoral dissertation
OFDM under Oscillator Phase Noise: Contributions to Analysis and Estimation Methods

Publisher School of Electrical Engineering

Unit Department of Signal Processing and Acoustics

Series Aalto University publication series DOCTORAL DISSERTATIONS 105/2016

Field of research Signal Processing for Communications

Manuscript submitted 19 November 2015 **Date of the defence** 28 June 2016

Permission to publish granted (date) 13 February 2016 **Language** English

☐ **Monograph** ☒ **Article dissertation** ☐ **Essay dissertation**

Abstract

Most modern transmitters and receivers involve an analog front-end unit and a digital back-end unit. The digital back-end is responsible for information processing which involves the following: redundancy removal from information; information representation; improving information resilience; and information correction. The analog front-end is responsible for information transmission and reception. The information processing algorithms developed and implemented in the digital back-end assume a linear and noiseless analog front-end which, in reality, is not the case. This renders some of the information processing algorithms to be less effective in practice. The focus of this thesis is on *orthogonal frequency-division multiplexing* (OFDM) systems under the influence of oscillator *phase noise*. OFDM is an efficient information representation technique used in many communication systems. On the other hand, phase noise is one type of undesired noise that occurs in the oscillator device used in the analog front-end. It arises due to the imperfect task of frequency conversion, performed by the oscillator device, between baseband and radio frequency.

This thesis contributes to the areas of analysis and estimation in OFDM systems under the influence of oscillator phase noise. With regard to analysis, this thesis contributes by deriving the channel capacity assuming a Gaussian input alphabet. The aim here is to show both quantitatively and qualitatively the degradation in performance of the OFDM system in the presence of phase noise. The analysis is conducted for phase noise processes that occur in both free-running and phase-locked loop based oscillators and also extended to include the effect of carrier frequency offset. With regard to estimation, two new phase noise estimation algorithms are proposed in this thesis. In particular, these algorithms restrict the *search* space to a specific subset, where the desired phase noise parameter of interest is shown to lie. For example, in the first estimation method, possible *subspaces* in which the desired phase noise spectral vector may lie are used in the estimation step. In the second method, the *geometry* of the desired phase noise spectral vector is used in the estimation step. Specifically, this geometry corresponds to a *non-convex* set described by a set of quadratic forms that involve permutation matrices. By restricting the search space to this set, the accuracy of phase noise estimation can be improved.

Keywords OFDM, Phase Noise, 4G, 5G, Estimation, Optimization, S-procedure

ISBN (printed) 978-952-60-6838-1 **ISBN (pdf)** 978-952-60-6839-8

ISSN-L 1799-4934 **ISSN (printed)** 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Helsinki **Location of printing** Helsinki **Year** 2016

Pages 207 **urn** http://urn.fi/URN:ISBN:978-952-60-6839-8

“What an astonishing thing a book is.
It’s a flat object, made from a tree with flexible parts, on which are imprinted lots of funny dark squiggles.
But one glance at it and you are inside the mind of another person maybe somebody dead for thousands of years.
Across the millennia, an author is speaking clearly and silently inside your head. Directly to you.
Writing is perhaps the greatest of human inventions.
Binding together people who never knew each other. Citizens of distant epochs.
Books break the shackles of time.
A book is proof that humans are capable of working magic.”

Carl Sagan

Preface

The research work pertaining to this doctoral thesis was conducted at the Department of Signal Processing and Acoustics which is a research and faculty unit in the School of Electrical Engineering, Aalto University. The research work was supervised and supported by Prof. Risto Wichman who leads the wireless signal processing research group. In addition, the Graduate School in Electronics, Telecommunications and Automation (GETA) provided the required funding for the research during the years 2011 and 2012. The Nokia Foundation is also acknowledged here for providing support in promoting the research ideas presented in this thesis.

At this point, I would first like to thank Prof. Risto Wichman for his unwavering support during these research years. It is highly encouraging for a doctoral candidate to know that he has his supervisor's faith and trust. Secondly, I would like to thank Dr. Taneli Riihonen and Dr. Stefan Werner who have been the pillars of my research work. You have trusted, guided, and shaped all of my ideas and judgments on our research topics. This research work would not have been possible without your support. Thank you for caring, being supportive and so easily approachable. I thank Prof. Mikko Valkama and Dr. Nikolay N. Tchamov for their valuable insight and feedback, particularly, on phase noise modeling. My gratitude also goes to Prof. Sergiy Vorobyov for discussions on optimization theory which helped cleared some of the haze in my understanding of the topic.

I thank my preliminary examiners, Assistant Prof. Hani Mehrpouyan and Dr. Prathapasinghe Dharmawansa, for agreeing to review this thesis. Your rigorous feedback has greatly improved the quality of this manuscript. I also thank Prof. Bhaskar D. Rao and Assistant Prof. Hani Mehrpouyan for willing to be my opponents. I am much humbled at this gesture and, hopefully, I can provide for some interesting discussions related to our

research. A special word of appreciation goes to Prof. Visa Koivunen for providing and facilitating the opportunity to have Prof. Bhaskar D. Rao as my opponent. Thank you.

My gratitude also goes to my colleagues who have given me many pleasant memories. Hopefully, I have done the same for you too. I have enjoyed spending time with you, having lunches together, our travel trips, our games of badminton in the early days with some strokes of cricket, bowling, floor ball and curling also. I vividly remember the Karaoke nights in Ruka, the midsummer and Christmas dinners where I have had the occasion to know you better with the aid of a beer glass. My gratitude goes out to all my dear friends, especially here in Finland. I have had so many wonderful experiences with you, and I am always in awe of your talent, your diverse ways of thinking about life and in conducting it.

No amount of appreciation and gratitude can equalize the love and support of ones parents and siblings. Thank you Mama, Dada, Tina and Pratap for always looking after and being there for me.

Espoo, May 11, 2016,

Pramod Jacob Mathecken

Contents

List of Publications

This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.

I P. Mathecken, T. Riihonen, S. Werner and R. Wichman. Performance Analysis of OFDM with Wiener Phase Noise and Frequency Selective Fading Channel. *IEEE Transactions on Communications*, vol. 59, no. 5, pp. 1321–1331, May 2011.

II P. Mathecken, T. Riihonen, N. N. Tchamov, S. Werner, M. Valkama and R. Wichman. Characterization of OFDM Radio Link under PLL-based Oscillator Phase Noise and Multipath Fading Channel. *IEEE Transactions on Communications*, vol. 60, no. 6, pp. 1479–1485, June 2012.

III P. Mathecken, T. Riihonen, S. Werner and R. Wichman. Average Capacity of Rayleigh-fading OFDM Link with Wiener Phase Noise and Frequency Offset. In *Proc. 23rd IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 2353–2358, September 2012.

IV P. Mathecken, T. Riihonen, S. Werner and R. Wichman. Accurate Characterization and Compensation of Phase Noise in OFDM Receiver. In *Proc. 45th Asilomar Conference on Signals, Systems and Computers (ACSSC)*, pp. 1948–1952, November 2011.

V P. Mathecken, T. Riihonen, S. Werner and R. Wichman. Phase Noise Estimation in OFDM: Utilizing its Associated Spectral Geometry. *IEEE Transactions on Signal Processing*, vol. 64, no. 8, pp. 1999–2012, April 2016.

VI P. Mathecken, S. Werner, T. Riihonen, and R. Wichman. Subspace-based Phase Noise Estimation in OFDM Receivers. In *Proc. 40th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3227–3231, October 2014.

Author's Contribution

Publication I: "Performance Analysis of OFDM with Wiener Phase Noise and Frequency Selective Fading Channel"

The idea of using a probabilistic approach for performance analysis was formulated together with the co-authors. The author of this thesis contributed by deriving all technical results, performing the numerical simulations and writing the entire manuscript. The co-authors have provided valuable insight that aided the author in the research work as well as guidance and constant feedback on the overall structure, presentation and content of the publication.

Publication II: "Characterization of OFDM Radio Link under PLL-based Oscillator Phase Noise and Multipath Fading Channel"

This work stems from collaboration with the co-authors from Tampere University of Technology. The objective of this work was to extend the performance analysis originally done in Publication I for Wiener phase noise processes to PLL-based phase noise processes. The co-authors from Tampere University of Technology mainly contributed towards PLL-based phase noise modeling which includes simulation scripts that generate such processes. The author of this thesis contributed to this publication by deriving all technical results, performing the numerical simulations related to performance analysis and writing the entire manuscript. The co-authors provided guidance and constant feedback on the overall structure, presentation and content of the publication.

Publication III: "Average Capacity of Rayleigh-fading OFDM Link with Wiener Phase Noise and Frequency Offset"

This work extends the analysis of Publications I and II to include the effect of carrier frequency offset. The author of this thesis contributed to this publication by deriving all technical results, performing the numerical simulations and writing the entire manuscript. The co-authors provided guidance and constant feedback on the overall structure, presentation and content of the publication.

Publication IV: "Accurate Characterization and Compensation of Phase Noise in OFDM Receiver"

The idea for the publication was spawned by the framework used in deriving probability density functions of Publications I, II and III. The author of this thesis contributed to this publication by deriving all technical results, performing the numerical simulations and writing the entire manuscript. The co-authors provided guidance and constant feedback on the overall structure, presentation and content of the publication.

Publication V: "Phase Noise Estimation in OFDM: Utilizing its Associated Spectral Geometry"

The idea spawned from the *unit-norm* property that was observed by the author of this thesis for the desired phase noise spectral components. The co-authors observantly pointed out that its time-domain manifestation is the constant-magnitude property of the complex exponential phasor. This immediately pointed towards a deeper and beyond the unit-norm property of the spectral components. The author of this thesis established and derived this deeper connection which was then utilized to perform phase noise estimation. The author derived all the technical results, performed the numerical simulations and wrote the entire manuscript. The co-authors provided guidance and constant feedback on the overall structure, presentation and content of the publication.

Publication VI: "Subspace-based Phase Noise Estimation in OFDM Receivers"

The idea to use *subspace* information was mainly spawned by the work done in Publication V. The author of this thesis was responsible for developing the original ideas, deriving all the technical results, performing the numerical simulations and writing the entire manuscript. The co-authors provided guidance and constant feedback on the overall structure, presentation and content of the publication.

List of Abbreviations

Abbreviation	Expansion
4G	Fourth generation
5G	Fifth generation
ADC	Analog-to-digital converter
AM	Amplitude modulation
AWGN	Additive white Gaussian noise
BEP	Bit-error probability
BER	Bit-error rate
BPSK	Binary phase-shift keying
CDF	Cumulative distribution function
CP	Charge pump
CPE	Common phase error
DAC	Digital-to-analog converter
DFT	Discrete Fourier transform
DSP	Digital signal processor
EM	Expectation maximization
E-step	Expectation step
GPS	Global positioning system
ICI	Inter-carrier interference
IDFT	Inverse discrete Fourier transform
IEEE	Institute of Electrical and Electronics Engineers
IF	Intermediate frequency
ISI	Inter-symbol interference
LC	Linear constraint

LF	Loop filter
LMI	Linear matrix inequality
LNA	Low-noise amplifier
LTE	Long term evolution
LTI	Linear time-invariant
MAP	Maximum a-posteriori probability
MGF	Moment generating function
MIMO	Multiple-input multiple-output
ML	Maximum likelihood
MLE	ML estimate
MMSE	Minimum mean square error
M-step	Maximization step
MSE	Mean square error
OFDM	Orthogonal frequency division multiplexing
OFDMA	Orthogonal frequency division multiple access
PA	Power amplifier
PAPR	Peak-to-average power ratio
PARMA	Parallel auto-regressive moving average
PC-PPT	Piecewise-constant PPT
PDF	Probability density function
PFD	Phase frequency detector
PLL	Phase-locked loop
PM	Phase modulation
PNC	Phase noise constraints
PPT	Phase-noise-geometry-preserving transformation
PSD	Power spectral density
RF	Radio frequency
RO	Reference oscillator
SDP	Semi-definite program
SEP	Symbol-error probability
SER	Symbol-error rate
SINR	Signal-to-interference-plus-noise ratio

SNR	Signal-to-noise ratio
SSB	Single side-band
UNC	Unit-norm constraint
VCO	Voltage controlled oscillator
WLAN	Wireless local area network

List of Symbols

Symbol	Meaning
$\bar{\mathcal{C}}$	Capacity averaged over phase noise
$\bar{\bar{\mathcal{C}}}$	Capacity averaged over phase noise and channel
$E_1(\cdot)$	Exponential integral function of order 1
$\mathbf{F}$	$N_\mathrm{c} \times N_\mathrm{c}$ DFT matrix
$\mathbf{F}_\mathrm{t}$	$N_\mathrm{c} \times L$ DFT matrix
$\tilde{\mathbf{F}}$	$N \times N$ DFT matrix
F_s	Sampling frequency
f_c	Carrier frequency
f_off	Frequency offset
f_sub	Subcarrier spacing
f_3dB	3-dB bandwidth of oscillator PSD
f_Δ	Normalized frequency offset
$\mathcal{G}(\alpha, \beta)$	Gamma distribution with parameters α and β
$h[n]$	Baseband channel
H_k	N_c-point DFT of $h[n]$
$\mathbf{H}$	Diagonal matrix with H_k as diagonal elements
$\mathbf{I}_N$	$N \times N$ identity matrix
$\jmath$	Imaginary unit
$\mathcal{L}(\cdot)$	Log-likelihood function
N_c	Number of subcarriers
N_cp	Length of cyclic prefix

$\mathbf{P}_{\mathrm{r}}$	$N_{\mathrm{c}} \times N_{\mathrm{c}}$ orthogonal projection matrix
$\mathbf{r}$	$N_{\mathrm{c}} \times 1$ received information vector
$\mathbf{s}$	$N_{\mathrm{c}} \times 1$ transmitted information vector
$\mathbf{S}$	Diagonal matrix with s as the main diagonal
T_{s}	Sampling period
$\mathbf{X}^{\mathrm{T}}$	Transpose of matrix $\mathbf{X}$
$\mathbf{X}^{\dagger}$	Hermitian transpose of $\mathbf{X}$
$\mathbf{X}^{+}$	Pseudo-inverse of $\mathbf{X}$
$\mathrm{span}(\mathbf{X})$	Span of columns of $\mathbf{X}$
$\mathrm{rank}\,(\mathbf{X})$	Rank of $\mathbf{X}$
$\mathrm{diag}\,(\mathbf{x})$	Diagonal matrix with vector x as the main diagonal
$\emptyset$	Empty set
$\mathrm{cov}\,(\mathcal{X})$	Convex hull of set $\mathcal{X}$
Y	ICI power
$\boldsymbol{\delta}$	N_{c}-point DFT of $e^{\jmath\theta[n]}$
δ_k	kth component of $\boldsymbol{\delta}$
$\theta[n]$	Discrete-time phase noise
$\theta(t)$	Continuous-time phase noise
ρ	Ratio of f_{3dB} to f_{sub}
σ_{s}^2	Transmitter signal power
σ_{w}^2	Receiver noise power

1. Introduction

Most modern communication devices consist of two processing units: *an analog front-end*; and *a digital back-end*. The digital unit processes *information* like any modern computer would do today, although there are distinct differences. The main task of the analog unit is to embed/detach information on/from a *carrier* signal. This information-bearing carrier signal is essentially what is transmitted from and received to the communication device, which unfortunately, for us humans, gets corrupted by the medium (also known as, the channel) through which it travels. This inevitably leads to corruption of information. In modern communication systems, the corrupted information is processed by the digital unit with the aim of recovery. To further strengthen the recovery process, the digital unit efficiently processes the information before transmission so that it is robust to the propagation medium. Traditionally, the corrupted information was corrected by the analog unit, typically, in an all-analog-communication system. With the advent of digital processing, the functionality of information correction transitioned to the digital part.

Digital processing in modern communication systems involve the following steps: Redundancy removal from information; information representation; improving information resilience; and information correction. The major leaps in modern telecommunication systems have mainly come from each of these areas. For example, in the latest wireless communication standard known as 'long term evolution' (LTE), a modulation technique called 'orthogonal frequency division multiplexing' (OFDM) is used for information representation. This technique facilitates a simple channel correction scheme, thereby, reducing the complexity of the digital processing unit. This modulation technique is also a potential contender for future communication systems as well. Almost all digital processing techniques developed today are done with the aim of combating the effects

of the propagation medium, while the sole purpose of the analog unit is the transmission and reception of the information-bearing carrier signal. These techniques are developed under the assumption of an ideal analog unit that is linear and noise-free. Unfortunately, this is not the case, and the analog unit itself distorts the information-bearing signal, mainly due to component imperfections that make up the analog circuitry. This essentially renders some digital processing techniques developed to combat the effects of channel less useful as they do not take into account the non-idealities of the analog unit.

There are three major functionalities of the analog unit that contribute substantially to the analog non-idealities. These are: Power amplification; up/down frequency conversion; and high-rate sampling. Power amplifiers typically operate in the non-linear region for reasons of efficiency. However, this contributes to non-linear distortion of the information-bearing carrier signal. The up/down frequency conversion refers to the process of embedding information onto the high-frequency passband carrier signal. This high-frequency signal is generated using an *oscillator* device. Unfortunately, these devices do not produce ideal carrier signals of the prescribed frequency, rather, there is random fluctuation in the frequency. Typically, the fluctuation is described in terms of the carrier phase, and it is popularly known by the name of *phase noise*. The function of high-rate sampling is to transition between the analog and digital domains. These are implemented by devices known as analog-to-digital and digital-to-analog converters. These devices introduce *jitter noise* which is random variations in sampling instants and is closely related with phase noise.

There are two ways to solve the problem of non-linear and noisy analog units: The first route is to design effective analog circuits and use high-quality devices that make up the analog circuitry. This inevitably increases the cost of the analog unit which maybe justified depending upon what the circumstances are; The second route is to use digital signal processing algorithms to remove the non-idealities. These algorithms can then be implemented in the digital unit. The downside, however, is added delay in the system which may be tolerable. In reality, both approaches for solving the non-ideality problem of the analog unit are being pursued. Extensive research, in the field of circuit design, is on going to design low-powered, cost-effective and spectrally pure analog devices. Simultaneously, signal processing engineers are making use of the increasing prowess of the digital unit to develop effective compensation algorithms.

1.1 Motivation, Scope and Objectives of the Thesis

This thesis investigates the problem of having phase noise in communication systems that employ the OFDM modulation.

The progress of science and especially engineering is generally based on the philosophy '*What is not broken need not be fixed*'. In line with this philosophy, the first step is to ascertain how serious of a problem is phase noise for a communication system employing OFDM. Extensive analysis has been conducted in the past twenty years on the performance of OFDM in the presence of phase noise, and the resounding conclusion is that it causes a significant drop in performance. Most performance metrics used to evaluate the effect of phase noise on OFDM have been in terms of signal-to-interference-plus-noise ratios, bit-error rate or the probability of a bit error. Another significant performance metric that is typically used in evaluating the efficacy of a communication system is the data rate or, technically, *channel capacity*. Evaluating capacities of linear systems is a well researched problem with off-the-shelf solutions. However, for non-linear systems like the analog unit in a communication system, the problem is hard and not so straightforward. One of the motivations of this thesis is to fill this gap in knowledge, and the associated objective is to precisely quantify the capacity degradation of an OFDM radio link impaired by phase noise.

The acknowledgment of performance loss has led researchers to seek new phase noise estimation and compensation schemes for OFDM. The literature is abundant with very good phase noise estimation schemes, and new algorithms still keep rolling out even to this day. The second objective of this thesis is to develop new signal processing algorithms to estimate phase noise and then remove it from the OFDM signal. This thesis contributes to the area of phase noise estimation by first recognizing certain properties of phase noise which are then utilized during the estimation process. In fact, these properties are well known in the community. However, this thesis shows a different manifestation of this property and how that can be utilized mathematically rather than using an ad-hoc approach. Of course, the study discusses trade-offs in using a complicated mathematically rigorous approach and a less complex ad-hoc approach.

1.2 Contributions of the Thesis

The main contributions of this thesis are in the areas of analysis and estimation for an OFDM radio link impaired by phase noise. These are summarized below.

C1. Analytical and closed-form expressions of capacity for an OFDM radio link under the influence of phase noise are derived. The analysis is conducted for two types of phase noise processes which occur in free-running and in phase-locked loop (PLL) based oscillator devices. For free-running oscillators, the phase noise follows a Wiener process. For modeling phase noise in PLL devices, the popular 'linear-time-invariance in phase-domain' model is used. The capacity analysis is extended to also include the effect of carrier frequency offset. Through these analytical expressions, the objective is to quantify the degradation in channel capacity while, at the same time, to obtain some qualitative insight.

The realization of the aforementioned contributions is achieved by a set of other contributions which are chiefly in the area of statistics and deal with determining probability density functions. These contributions are summarized next.

- The probability density function (PDF) of a sum of correlated gamma random variables with the same alpha parameter is derived. The previous state-of-the-art result is under the assumption of full-rank normalized covariance matrix of these gamma random variables while this thesis extends the result to the general case of any rank. This PDF is then applied to determine the average channel capacity for an OFDM radio link impaired by phase noise.
- The PDF of a sum of correlated gamma and Gaussian random variables is derived. The resulting distribution has a form similar to the PDF of sum of correlated gamma random variables. This distribution is used in the evaluation of average capacity when the OFDM system is impaired by both phase noise and carrier frequency offset.
- The above result holds for a particular structure of the correlated gamma and Gaussian random variables. In this thesis, the result is extended to the generic case, where the resulting PDF can be decoupled in terms of two independent random variables: one follows a Gaussian distribution, while the second random variable has a distribution similar to that of a sum of correlated gamma random variables.

C2. Two novel phase noise estimation schemes are developed in this thesis. Both of these schemes make use of specific information about phase noise during estimation. For example, in the first method, *subspace-based* information is utilized in obtaining a phase noise estimate, where possible subspaces in which the desired quantity may lie are explored. The second method utilizes the geometry associated with phase noise in the context of OFDM. Specifically, we refer to the geometry of the *spectral components of the complex exponential of phase noise*. This geometry is described by a set of *non-convex quadratic forms* that involve permutation matrices. Equivalently, in the time domain, this property manifests itself as *unit-magnitude* time domain samples. This property is utilized during the estimation step, where we enforce the phase noise estimates to satisfy this property.

The set of contributions aiding realization of the above are:

– A new linear phase noise spectral model is presented for the purpose of dimensionality reduction. Dimensionality reduction eases the estimation process since only a few number of components, less than the total number of dimensions, are to be estimated. The complex exponential of phase noise are low-pass processes, thereby, their spectral content is limited to only a few low-frequency components. At the same time, the complex exponential of phase noise signal has a specific geometry which gets destroyed when performing dimensionality reduction. In this thesis, a novel linear model is developed that performs dimensionality reduction while, at the same time, preserving the geometry.

– The second contribution falls in the area of *optimization theory*. The task is to minimize a homogeneous quadratic function subject to non-convex quadratic equality constraints that involve permutation matrices. To solve this problem, this thesis uses the so-called *S-procedure* which was originally developed for *inequality* constraints. This thesis provides conditions for the S-procedure to be lossless for equality constraints. The S-procedure developed for equality constraints is then applied to solve the phase noise minimization problem. Specifically, by using the S-procedure, the minimization problem can be equivalently solved by solving a *convex dual problem* which has polynomial-time computational complexity.

1.3 Structure of the Thesis

This thesis is organized in three parts.

Chapter 2 is primarily about modeling. The necessary background and context, associated with this thesis, is first presented. Specifically, we introduce the *direct-conversion* transceiver architecture and discuss the various sources of noise that arise in the analog unit of this architecture type. We then discuss the principle of OFDM and, in particular, the relation to its *single-carrier* counterpart. The second part of the chapter is on modeling: The input-output relation for a transmitter receiver pair in the presence of phase noise is described; Well-known models on phase noise processes for free-running oscillators and phase-locked loop based devices are presented. The models presented in this chapter serve as the foundation for the rest of the chapters.

Chapter 3 is primarily about performance analysis of the OFDM radio link in the presence of phase noise. We first present some state-of-the-art methods that seek to precisely quantify the signal-to-noise ratio and the probability of bit and symbol errors. The second part of the chapter summarizes the contribution of this thesis which is to evaluate the channel capacity. The goal is to determine closed-form expressions of the capacity for an OFDM radio-link impaired by phase noise and frequency selective fading. The analysis is extended to also include the effect of carrier frequency offset.

Chapter 4 is about estimation in OFDM systems under the influence of phase noise. The main quantities to be estimated are: the channel; phase noise; and the transmitted symbols. We first present a generic classification of estimation methods which fall in the categories of either *isolated approaches* or *joint approaches* and then review some of the state-of-art methods which fall in either of them. Finally, the second part of this chapter summarizes two novel phase noise estimation schemes proposed in this thesis. Specifically, these methods rely on using information on *where* the desired phase noise dependent parameter may lie.

2. Modeling

2.1 Background

The following saying aptly summarizes the field of telecommunications and, in particular, wireless communications: *'The only thing constant in life is change'*. We have transitioned through three generations of mobile telecommunications technology and, only very recently, entered the fourth generation (4G) [1–3]. Research has already begun towards developing fifth generation (5G) communications technologies with the goal of a possible standard by the year 2020 [4]. This need for exploring and researching new communication techniques is mainly driven by an addictive thirst for exorbitant data rates which from the users perspective is mainly about the variety of services available at their disposal. It has been forecasted and fortunately recognized that the data rates offered by the current 4G communication technologies will fall short of the demand in the coming decade [5–7].

One of the most important requirements imposed on mobile telecommunication systems is *connectivity*, i.e., no matter where and when, the mobile user can always communicate. One way of satisfying this requirement is by having many mobile communication systems with different technologies coexisting together such that the mobile user can seamlessly roam between these systems depending upon the requirement. For example, today's commercially available 4G mobile smart phones support access to both UMTS and GSM networks which are 3G and 2G systems, respectively. They also come equipped with Bluetooth, FM radio, the ever useful global positioning system (GPS) and finally, WiFi. All of these communication systems typically operate in different frequency bands and, in general, employ different communication techniques. The downside of

having together different communication systems is that it places different requirements on the *hardware* of the mobile radio terminal [8]. With the advent of 4G and upcoming 5G communication systems, extensive research is on going in the field of mobile radio architectures that can handle multi-standard, multi-mode and multi-band operations [9–11].

One of the most common transceiver hardware architecture in use today is the *direct-conversion* architecture [9, 12, 13]. A brief survey in [9] reveals that more than fifty percent of the transceiver architecture literature was devoted to this type. This architecture is mainly popular for its *integrability* and *low power consumption* which makes it very amenable for multi-standard, multi-mode and multi-band operations. High integrability and low power consumption essentially imply compact mobile terminals with long battery life. Of course, the direct-conversion architecture is one among many potential architectures for multi-standard transceivers. A good overview of these different architectures can be found in [9,14–16].

The hardware of a typical mobile radio device consists of two parts: an *analog part* and a *digital part*. The analog part, from its name, is made up of electronics that typically process continuous-time signals at radio frequencies (RF). This analog part is typically referred to in the literature as the *RF front-end*. The digital part deals with discrete-time signals whose main functionality is information processing and representation; The objective of the analog part is for signal transmission and reception. Unfortunately, the devices used in the make up of the RF front-end are non-ideal, whereby they either introduce undesired noise or cause distortion to the transmitted/received signal. These non-idealities are typically referred to as *RF-impairments* [17].

This chapter is structured in two parts: The first part of the chapter concerns with the problems encountered at the RF front-end of the mobile terminal. Using the direct-conversion architecture as an example, we briefly review the main functionalities and associated problems of the analog unit with a specific focus on *frequency conversion* achieved by means of the *oscillator* device that introduces phase noise. We then discuss orthogonal frequency-division multiplexing (OFDM) which can be interpreted as an information representation technique that is handled in the digital part of the mobile hardware. OFDM is used in 4G communication technologies and is a potential candidate for 5G systems as well [18]. The second part of this chapter is about modeling. Specifically, we model the communication link employing OFDM that includes phase noise at both

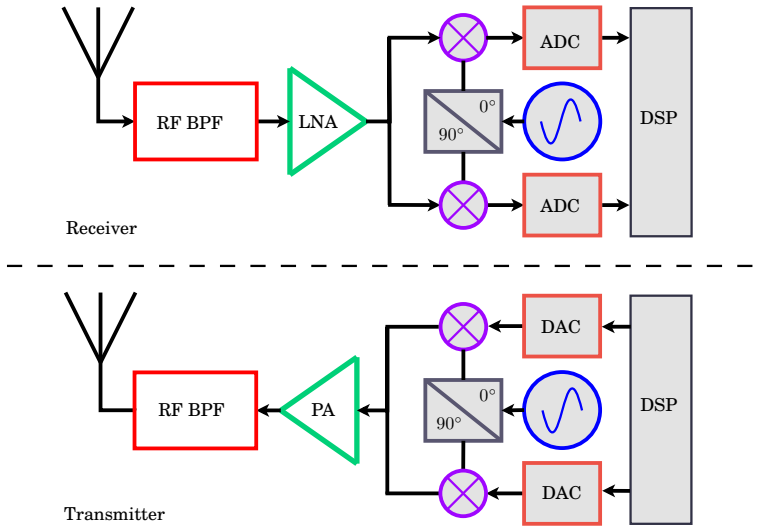


Figure 2.1. Constituents of a direct-conversion transceiver.

the transmitter and receiver ends. This communication model serves as the foundation for Chapters 3 and 4. The last part of this chapter is dedicated to phase noise modeling. The aim here is to briefly review some of the standard phase noise models that are also used in this thesis.

2.2 Direct-Conversion Transceiver

A direct-conversion transceiver is shown in Fig. 2.1. It does frequency conversion between RF and baseband frequency *directly* thereby avoiding an intermediate frequency (IF) stage typically found in other architectures like the *super-heterodyne and low-IF architectures* [9]. The absence of the IF-stage allows for on-chip integration which in turn results in low power consumption and compact terminals. This is because filters for image rejection and channel selection that accompany the IF-stage typically are implemented with passive components for improved performance.

In Fig. 2.1, the digital part of the direct-conversion transceiver is the digital signal processor (DSP), where all the baseband digital algorithms are implemented. The RF front-end or analog domain consists of three main functionalities: analog-to-digital/digital-to-analog conversion (ADC/DAC), frequency conversion between baseband and RF; and, finally, power amplification. The amplification of signals is performed by the low noise amplifier (LNA) and power amplifier (PA) in the receiver and transmitter, respectively. The frequency conversion is performed by the mixers (shown in

purple color) with aid of an *oscillator* (shown in blue color) which provides the reference carrier signal of the prescribed frequency. In practice, the oscillator is never used in isolation, and it is typically used in a feedback mechanism like in a phase-locked loop. Without any loss in generality, we assume an oscillator device feeding the mixers.

The analog front-end of the direct-conversion transceiver is, unfortunately, not ideal and the devices that perform the RF functionalities come with their own problems. Let us briefly discuss some of these issues.

2.2.1 Phase Noise in Oscillators

Oscillators are devices that produce periodic signals which are useful primarily because they help keep track of time. In this context, they are typically referred to as clocking signals. One cannot think of any digital device without any clocking signal involved. In the context of wireless communications, periodic *cosinusoidal* signals, produced by an oscillator, are information carriers. We build communication systems wherein information is typically represented in baseband and, for the purpose of transmission, we *embed* this information on a high-frequency cosinusoidal signal.

Figure 2.2 shows this information embedding process. The oscillator (shown in blue color) outputs a cosinusoidal signal of frequency f_{c}. Its corresponding frequency spectrum is shown below which consists of two delta functions centered at f_{c} and $-f_{\mathrm{c}}$ respectively. This signal is then mixed, using a mixer (shown in purple), with the baseband information-bearing signal whose frequency spectrum is centered around the zero frequency (shown in green). Mathematically, the mixing process is essentially a multiplication between the inputs to the mixer which, equivalently in the frequency domain, is the convolution operation. Thus, the result of the mixing operation is the translation of baseband frequency content to the high-frequency passband region centered around f_{c}, as seen in the figure. The information embedding process shown in Fig. 2.2 is idealistic and, in reality, the embedding process is flawed mainly due to imperfections of the oscillator and mixer devices. We shall assume herein that we have an ideal mixer and focus our attention only on the oscillator device.

Any practical oscillator does not generate pure cosinusoidal signals with spectrum as shown in Fig. 2.2. In practice, there is spectral spreading around the carrier frequency f_{c}, as shown in Fig. 2.3. This spectral spread is mainly attributed to two physical quantities of interest: These are the so-called *phase noise* and *amplitude noise*. We shall assume that ampli-

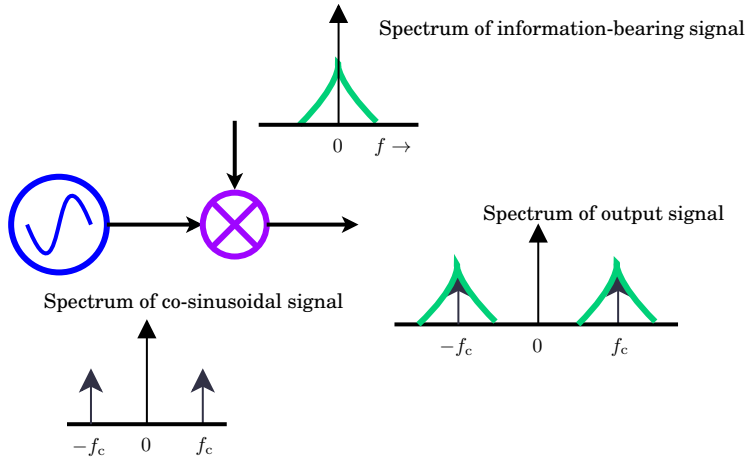


Figure 2.2. Information embedding process through mixing. The carrier frequency is denoted by f_c.

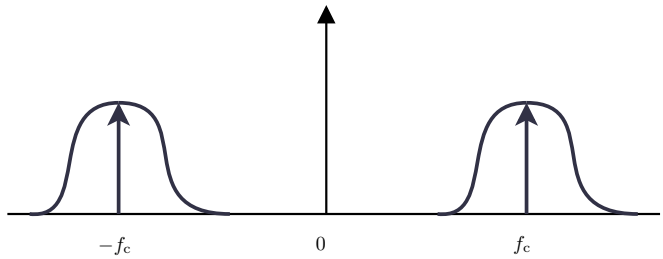


Figure 2.3. Illustration of spectral spreading in an oscillator. The carrier frequency is denoted by f_c.

tude noise is kept tolerable and that phase noise is the sole contributor to the spectral spread of the oscillator frequency spectrum [19].

Phase noise is the random perturbation in the phase of the cosinusoidal signals. It arises due to inherent noise present in any physical device such as the oscillator but it can be kept tolerable by proper choice of oscillator design [19]. For small spectral spreads, phase noise is slowly varying, while larger spectral spreads result in fast-varying phase noise processes. The spectral spread of the oscillator signal essentially distorts the information bearing signal because the information bearing signal spectrum is convolved with the spectrum shown in Fig. 2.3 which results in a distorted spectrum. In Section 2.3, we shall see how this distortion takes place in the case of OFDM.

2.2.2 Jitter in ADC and DAC

ADCs and DACs act as the bridge that connect the analog and digital domains of a communication transceiver. For the ADC, the goal is to convert the incoming analog signal into a digital signal which is then processed by the digital signal processor; The reverse operation is performed by the DAC. Today's and next-generation communication systems place phenomenal challenges on the ADC and DAC, where the mobile transceiver is expected to work in frequency ranges well over a few gigahertz. One of these challenges is high sampling rates which are in proportion to the bandwidth of the communication signals. Especially, LTE signals are wideband and the future 5G communication signals will also see large bandwidths [2, 20]. Another challenge is high resolution of the ADC which is difficult to obtain at high sampling rates and the last important design factor is power dissipation. Recent trends in ADC and DAC design have mainly focused towards achieving low power dissipation which eventually allows for efficient system-on-chip integration. A comprehensive overview of ADCs and the trade-offs that exist between the aforementioned design factors can be found in [21–24].

Hindering the challenges of high sampling rate and resolution, while lowering the power dissipation, are the various noise sources that creep into the ADC and DAC. The typical noise sources are jitter noise, quantization noise, thermal noise and non-linearities [21]. Jitter noise, especially, becomes prominent at high sampling rates and can significantly alter the spectrum of the digitized signal. It refers to random fluctuations in the sampling instants used to discretize the incoming analog signal. This noise typically occurs in the *sample-and-hold* circuitry that is responsible for the discretization process. It also occurs due to phase noise in the oscillators that supply the clocking signal to the ADC. This type of fluctuation is typically referred to as *clock jitter*. The overall jitter noise is a combination of both these types of jitter, and the end result is randomness in the sampling instants.

The research community has devoted efforts to study the effects of jitter noise on the resulting output signals. Many different characterizations exist with each tackling a particular aspect depending upon the goal [25]. For example, in [26] and [27, 28], the spectrum of the output signal for an ADC and DAC corrupted by timing jitter are derived, respectively. In Fig. 2.4, we show an illustration of the output spectrum of an ideal DAC

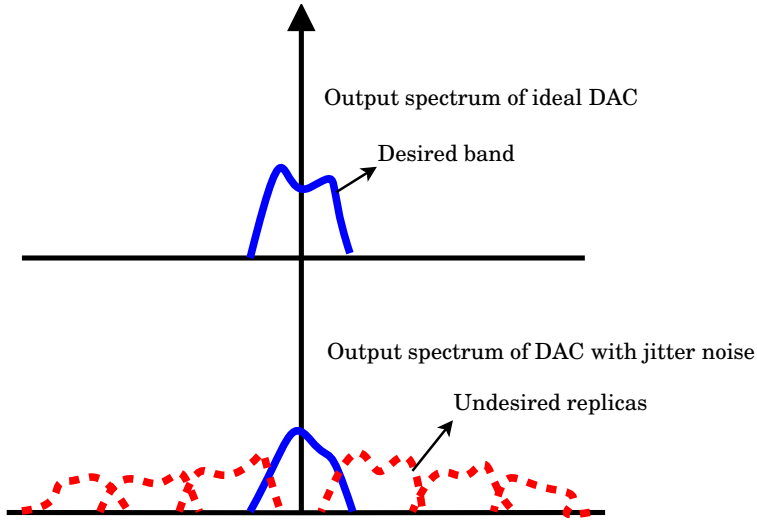


Figure 2.4. Illustration of the effect of jitter noise.

and a DAC corrupted with timing jitter. The effect of jitter is to *alter the shape* of the desired spectrum while at the same time generate *copies* of desired spectrum. A similar effect is seen in the spectrum of the output signal for an ADC. The reader may refer to [26]; The result, however, is restricted in the sense that, although the timing jitter is assumed random, a certain periodic structure for the timing jitter is assumed.

With respect to OFDM, studies have recently been rolling out to analyze and compensate the effects of jitter noise [29–33]. Specifically, it has been shown that jitter noise has two effects on OFDM: The first is an additive noise contribution; and the second is a multiplicative noise contribution very similar to that of phase noise [30, 32]. By treating the additive noise as extra receiver noise, any phase noise estimation and compensation algorithm can then be applied to remove the multiplicative effect of jitter noise. See, for example, [32].

2.2.3 IQ-imbalance in Modulation and Demodulation

In direct-conversion transceivers, frequency conversion between RF and baseband frequency is implemented using a *quadrature* architecture [12]. For example, in Fig. 2.1, the incoming RF signal is split into two paths, namely *I-branch* and *Q-branch*, where each path is mixed with local oscillator signals that have a 90° phase difference between them. However, this is an idealistic scenario, and in practice, the phase difference is never exactly 90°, thereby, resulting in some correlation between the local oscil-

lator signals that feed the mixers. This mismatch in phase difference is typically referred to as *IQ-imbalance* or *IQ-mismatch*. As an example, the effect of IQ-mismatch on the signal of interest is shown in Fig. 2.5. In the ideal case of perfect 90° phase difference, the desired baseband spectrum (shown in blue) sits neatly around the zero frequency. However, in the more practical case of non-zero IQ-imbalance, the desired spectrum experiences interference from its *negative half*. A good mathematical treatment of the effect shown in Fig. 2.5 can be found in [34].

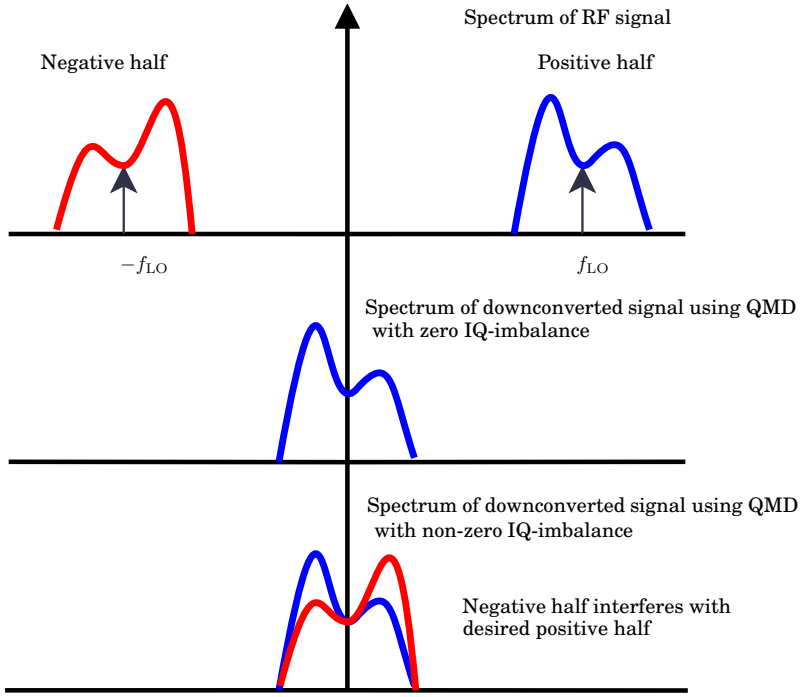


Figure 2.5. Illustration of the effect of IQ-imbalance.

The resulting interference and its level of impact depend on the amount of IQ-mismatch present in the hardware radio transceiver and also on the type of baseband signal used. It is well known in the scientific community that OFDM signals are sensitive to IQ-imbalance, and they yield poor performance in the presence of this mismatch [35]. Numerous studies have been undertaken to characterize this mismatch, and to develop effective signal processing algorithms to undo the effect of IQ-imbalance. See, for example, [36, 37] for a comprehensive treatment on the subject.

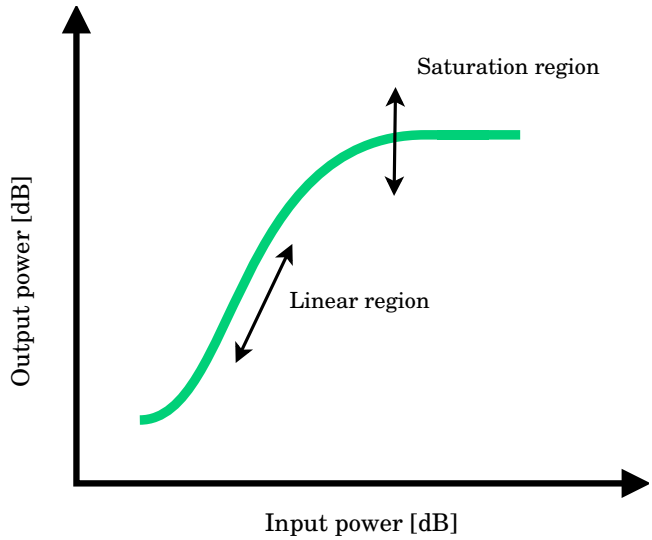


Figure 2.6. Illustration of the input-output characteristic of a power amplifier.

2.2.4 Non-linearity of Power Amplifiers

An ideal power amplifier would be linear, i.e., its output is a linearly amplified version of its input signal. Unfortunately, this is true only for a certain power region of the input signal, and if the input signal power exceeds this region then the output signal power gets saturated to a particular value, thereby, resulting in phenomena which are broadly classified as in-band and out-of-band distortion [38, 39]. The out-of-band distortion essentially leads to interference in neighboring channels.

Maintaining linearity of the power amplifier is a major design criteria especially for the transmitter. A typical power amplifier input-output characteristic is shown in Fig. 2.6. From the figure, we see that only a section of the input-output curve corresponds to a linear region and beyond this region, the signal power is saturated or compressed to a particular value. Especially, OFDM signals are known to have high *peak-to-average power ratio* (PAPR) which can result in driving the power amplifier to the saturation region which will result in signal distortion [40]. On the other hand, there is an inverse relationship between *power amplifier efficiency* and PAPR [41, 42]. Thus, we see that OFDM signals result in poor power amplifier efficiency which implies higher heat dissipation and, hence, poor battery life for the mobile terminal.

There are two possible ways to solve this problem: In the first method, the power amplifier is operated by employing high *back-off* which is a

measure of the region in which the PA operates. A high back-off implies that the power amplifier is operating in the linear region and, hence, the non-linear distortion due to saturation is avoided. In the second method, the power efficiency is improved by using low back-off while the associated non-linearity problem is solved by using *linearization techniques*. This is done by employing a *pre-distorter* such that the combined input-output characteristic of the pre-distorter and PA results in a linearized response [42]. Designing a pre-distorter requires accurate modeling of the PA with well defined models already available in the literature [43, 44]. Development of digital pre-distorters is still an active area of research. A good overview and extensive literature on the subject can be found in [38].

2.3 OFDM

We now turn our focus towards OFDM which is by far one of the most popular modulation schemes in use today. It is essentially an information representation technique with the principle aim of facilitating simple baseband transmitter and receiver structures. Specifically, in the receiver, the necessary functionality of channel equalization for an OFDM signal is implemented by trivial one-tap filters unlike in *single-carrier systems*. OFDM falls under the class of *multi-carrier* signals and comparison is typically done with its single-carrier counter part which requires involved filtering for channel equalization [45, 46]. We now illustrate the basic principle behind OFDM and where it differs from its single-carrier counterpart.

2.3.1 Principles of OFDM

In a single-carrier system, a set of N_{c} symbols, denoted by s_k, are transmitted using a sinc waveform denoted by $s(t)$. In practice, sinc waveforms are never used as they extend infinitely in time and, thus, some form of truncation is always done to the waveform. The symbols are transmitted at a rate $W = 1/T$, where T denotes the symbol duration. An illustration of a single-carrier signal is shown in Fig. 2.7, where three symbols are transmitted. As seen in Fig. 2.7, the symbols s_k are multiplied with time-shifted versions of the sinc function, however, the symbols are still recoverable because, at the zero-crossings, the sinc waveforms do not interfere with each other. The zero-crossing points are also known as *inter-symbol*

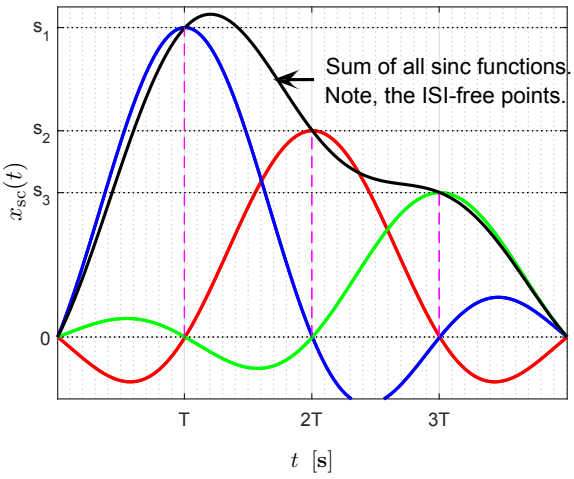


Figure 2.7. Single carrier waveform.

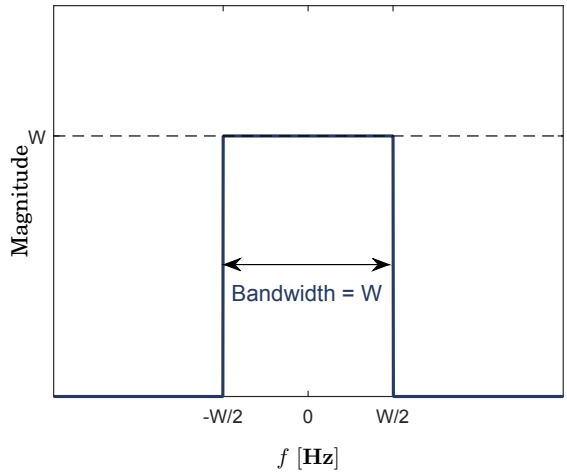


Figure 2.8. Frequency response of the sinc waveform.

interference (ISI) free instances and is shown by the pink dotted lines in Fig. 2.7. The corresponding frequency response of the sinc waveform is shown in Fig. 2.8. Thus, for the transmission of each symbol s_k, the total bandwidth used is equal to W Hz.

The drawback of using the single-carrier signal is that when passed through the channel, the necessary task of channel equalization has high complexity. This is mainly because the transmitted symbols s_k are represented in the time domain, and after the signal is convolved with the channel, at the ISI free instances, there is non-zero contribution from the

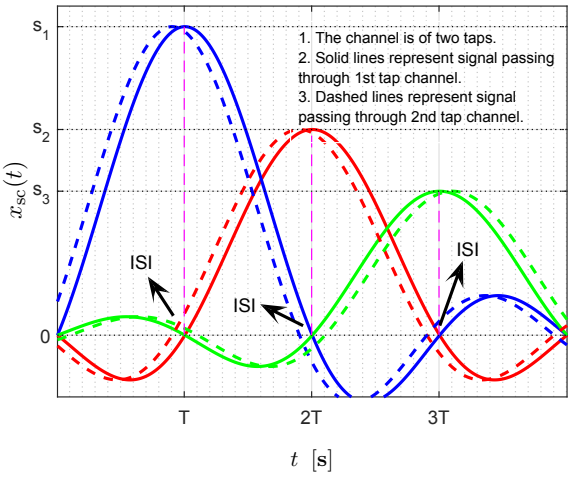


Figure 2.9. Single carrier waveform when passed through a two tap channel.

other sinc waveforms. This effect is shown in Fig. 2.9, where the channel in consideration is a simple two-tap channel. At the ISI-free instances, the desired symbol s_k experiences distortion from the signal passed through the second tap channel (represented by the dashed line) and also experiences additive interference from neighboring symbols.

In an OFDM signal, the transmitted symbols are represented in the frequency domain using sinc waveforms as shown in Fig. 2.10. The difference in comparison to its single-carrier counterpart is that the assigned bandwidth for each symbol s_k is compressed by a factor of N_c, where N_c is the number of transmitted symbols. In Fig. 2.10, $N_\mathrm{c} = 3$. As with the single-carrier case, the symbols are still recoverable because of the ISI-free points in the frequency-domain. Thus, all we need to do at the receiver side is to sample at these points in the frequency domain.

Embedding the information symbols s_k on sinc functions, in the frequency domain, facilitates a simple channel equalizer. Since, the signal is convolved with a time-domain channel, equivalently, in the frequency domain, the OFDM frequency response is *multiplied* with the channel frequency response. This effect is shown in Fig. 2.11. As seen in the figure, even after the multiplication operation, at the ISI-free points, there is no interference from neighboring symbols. Thus, at the receiver side, after sampling at the zero crossings in frequency domain, we can just divide by the estimate of channel frequency response to obtain estimates of the transmitted symbols. Essentially, the channel equalizer is a one-tap filter.

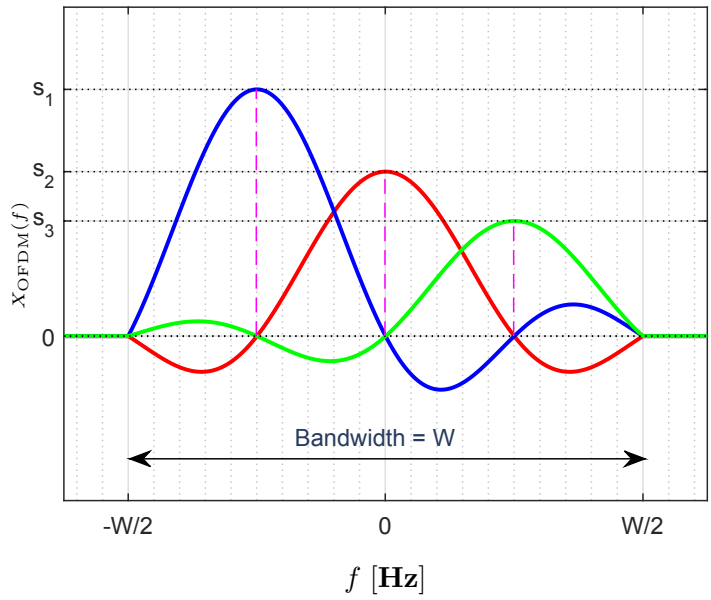


Figure 2.10. Frequency response of OFDM signal.

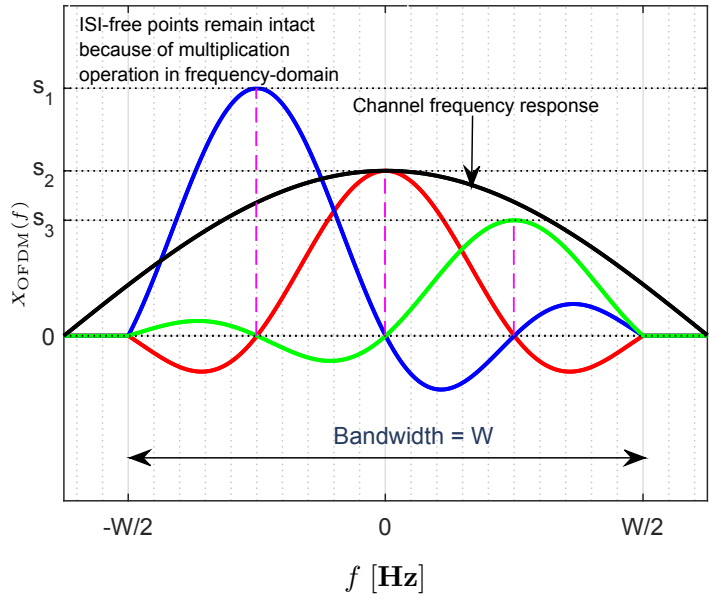


Figure 2.11. Illustration of an OFDM signal when passed through channel.

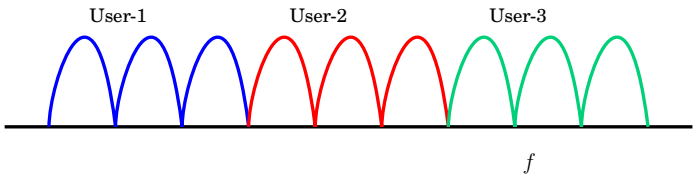


Figure 2.12. Subcarrier assignment of users in OFDMA.

2.3.2 OFDM in LTE-Advanced

In this section, we briefly touch upon some practical applications of OFDM. OFDM has been used in many communication systems, for example, in WLAN (wireless local area networks) systems and in DVB and DAB systems which are digital broadcasting systems for transmission of high quality video and audio. OFDM is also used in the latest wireless system known as LTE-Advanced which also encompasses newer and enhanced technologies like carrier aggregation, co-ordinated multi-point transmission, relaying, MIMO techniques and heterogeneous networks [2].

In a practical scenario, many mobile users compete for the same and limited radio resources offered by the wireless network. Thus, the available bandwidth must be shared between mobile users in a certain manner. In LTE-Advanced, the following channel bandwidths have been specified: 1.4 MHz, 3 MHz, 5 MHz, 10 MHz and 20 MHz [47]. This offers to the mobile operator a certain degree of flexibility. A simple illustration of bandwidth sharing between different users is shown in Fig. 2.12. In the figure, different users are allocated a subset of OFDM subcarriers. Such a scheme for example is used in the down-link of LTE-Advanced. This method of *multiplexing* users using OFDM is popularly known by *orthogonal frequency division multiple access* (OFDMA). Of course, various choices exist for mapping of subcarriers to users. For example, the subcarrier mapping to users in Fig. 2.12, is contiguous in nature. In a distributed mapping, users are assigned to non-contiguous subcarriers.

To enable efficient distribution and scheduling of physical layer resources among mobile users, a 'standard unit' of resource needs to be defined. Such a specification of a standard unit in LTE-Advanced is shown in Fig. 2.13. A 'physical resource block' in LTE-Advanced is defined as a group of twelve subcarriers for one slot of an LTE frame. Typically, one slot of length 0.5 milliseconds consists of seven OFDM symbols but can also contain six OFDM symbols [48]. In Fig. 2.13, the group of blocks

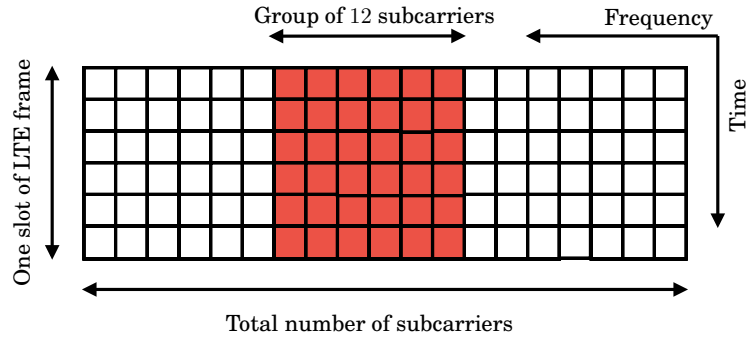


Figure 2.13. A physical resource block (all blocks in red color) in LTE-Advanced. Each slot of an LTE frame typically consists of seven OFDM symbols.

shown in red color constitute one physical resource block. The base station is primarily responsible for scheduling the mobile users by proper allocation of physical resource blocks. For example, each user may be allocated dedicated non-overlapping physical resource blocks. On the other hand, a single physical resource block may be shared between users.

We end this section with a brief comment on the future of OFDM, especially in the upcoming 5G wireless systems. It has been recognized that OFDM is not the ideal waveform, and it has its drawbacks. For example, in LTE-Advanced itself, in the up-link a modified version of OFDM, also called single-carrier frequency division multiplexing, is used. Such a modification is used because OFDM signals have high PAPR which leads to higher power consumption. Such a situation may not be tolerable for mobile devices, especially in the smart-phone business which is a fiercely contested market. This particular drawback and others are encouraging researchers to seek new multi-carrier waveforms and also to modify the existing OFDM waveform [49]. For example, *constant envelope* OFDM is one such variant of OFDM that seeks to alleviate the high PAPR of conventional OFDM [50].

2.3.3 OFDM System Model with Phase Noise

In this section, the mathematical formulation of a communication link impaired by both transmit and receive phase noise is presented. The communication link employs the OFDM modulation scheme. OFDM is essentially an information representation method, wherein information symbols are packed using sinc functions in the frequency domain. The process of embedding information symbols using OFDM is implemented

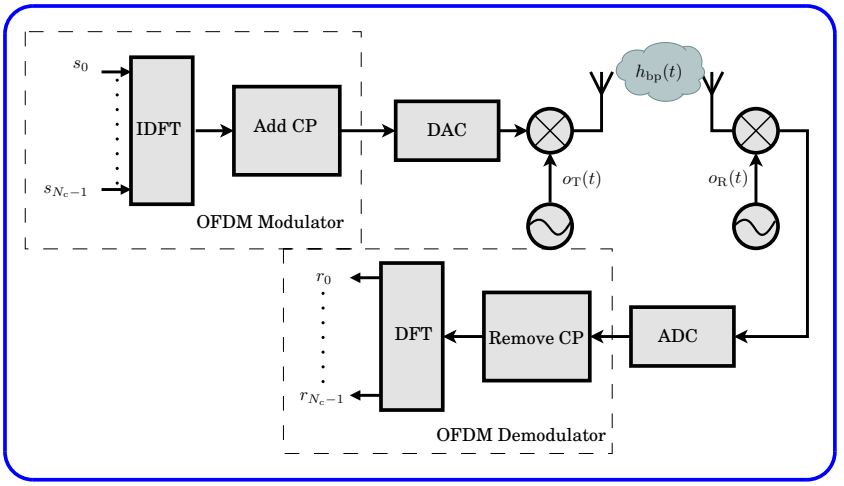


Figure 2.14. OFDM system model.

in the DSP of the direct-conversion transceiver of Fig. 2.1, while the analog front-end is responsible for the transmission of the baseband OFDM signal. In this section, a brief overview of the functionality of an OFDM modulator and demodulator are presented. We assume that the oscillator device of the analog front-end is the only source of noise; while the associated problems from other devices are either negligible in nature because of high quality devices used or already compensated for. Our aim here is to present the input-output relation that takes into account phase noise at the transmitter and receiver, respectively.

Figure 2.14 shows a communication link with an OFDM modulator and demodulator at the transmitter and receiver, respectively. The $N_c \times 1$ vector $\mathbf{s} = [s_0\ s_1 \dots s_{N_c-1}]^T$ represents the information to be transmitted using OFDM. This is done by first performing an inverse discrete Fourier transform (IDFT) on the sequence s_k. With such an operation, the symbols s_k are mapped to orthogonal and windowed exponential sinusoidal signals whose frequency response is given by the sinc function as seen in Fig. 2.10. The result of the IDFT operation is a time-domain signal. The *cyclic prefix* block appends a small amount of redundancy to the OFDM signal. This is done to counter the effect of ISI between OFDM blocks. The amount of redundancy, quantified by the length of the cyclic prefix denoted by N_{cp}, should typically be greater than the channel impulse response length. The resulting discrete signal is then transformed to the analog domain by the DAC and up-converted to the RF carrier frequency, denoted by f_c, by means of the mixer device and oscillator. The oscillator

delivers the carrier signal $o_{\mathrm{T}}(t) = e^{\jmath(2\pi f_c t + \theta_{\mathrm{T}}(t))}$ whose phase varies randomly because of the phase noise $\theta_{\mathrm{T}}(t)$. The signal is then transmitted through a bandpass channel denoted by $h_{\mathrm{bp}}(t)$.

At the receiver side, the received RF signal is corrupted by additive receiver noise (not shown in figure) and down-converted to baseband frequency by mixing with the oscillator signal $o_{\mathrm{R}}(t) = e^{-\jmath(2\pi f_c t - \theta_{\mathrm{R}}(t))}$, where $\theta_{\mathrm{R}}(t)$ denotes the receiver phase noise. The inverse operations of the transmitter side are performed, i.e., analog-to-digital conversion, removal of cyclic prefix and, finally a discrete Fourier transform (DFT) operation to obtain the received symbol vector $\mathbf{r} = [r_0\ r_1 \dots r_{N_{\mathrm{c}}-1}]^{\mathrm{T}}$. The expression relating $\mathbf{r}$ and $\mathbf{s}$, based on the functional blocks in Fig. 2.14, can be easily derived and is given by [51, 52]

$$\mathbf{r} = \mathbf{V}_{\mathrm{R}}\mathbf{H}\mathbf{V}_{\mathrm{T}}\mathbf{s} + \mathbf{w}, \tag{2.1}$$

where $\mathbf{H}$ is a diagonal matrix composed of elements $\{H_k\}_{k=0}^{N_{\mathrm{c}}-1}$ which are the DFT of $h[n]$, i.e.,

$$H_k = \sum_{n=0}^{N_{\mathrm{c}}-1} h[n] e^{-\jmath(2\pi kn)/N_{\mathrm{c}}}, k = 0, 1, \dots, N_{\mathrm{c}} - 1. \tag{2.2}$$

The quantity $h[n]$ is the discretized version of the low-pass equivalent of the bandpass channel $h_{\mathrm{bp}}(t)$ [53]. The vector $\mathbf{w}$ denotes the white Gaussian receiver noise with diagonal covariance matrix whose diagonal values are equal to σ_{w}^2. The unitary matrix $\mathbf{V}_{\mathrm{x}}$, $\mathrm{x} \in \{\mathrm{T},\ \mathrm{R}\}$, is column-wise circulant with the first column vector $\boldsymbol{\delta}^{\mathrm{x}}$ whose elements are given by

$$\delta_k^{\mathrm{x}} = \sum_{n=0}^{N_{\mathrm{c}}-1} \frac{e^{\jmath\theta_{\mathrm{x}}[n]}}{N_c} e^{-\jmath(2\pi kn)/N_{\mathrm{c}}}, k = 0, 1, \dots, N_{\mathrm{c}} - 1, \tag{2.3}$$

where $\theta_{\mathrm{x}}[n]$ is the discretized version of $\theta_{\mathrm{x}}(t)$ which is the continuous-time phase noise.

Equation 2.3 is nothing but the DFT of the complex exponential of the phase noise. The reader would come to appreciate (2.1), especially, when observing its association with that of Fig. 2.14. From Fig. 2.14, at the transmitter side, after the DAC operation, the time-domain signal gets multiplied with oscillator signal in the mixer device. This amounts to a convolution operation in the frequency domain, and the result, in (2.1), is represented by the quantity $\mathbf{V}_{\mathrm{T}}\mathbf{s}$, where $\mathbf{V}_{\mathrm{T}}$ is the circular convolution matrix. The signal then goes through the channel which in the frequency domain amounts to multiplication and, thus, we have the term $\mathbf{H}\mathbf{V}_{\mathrm{T}}\mathbf{s}$, where $\mathbf{H}$ is a diagonal matrix. At the receiver side, in the time domain,

another multiplication operation between the received signal and the oscillator signal which in frequency-domain is represented by a convolution operation and, hence, the result is given by $\mathbf{V}_{\mathrm{R}}\mathbf{H}\mathbf{V}_{\mathrm{T}}\mathbf{s}$. At this point, we caution the reader that the aforementioned description is not a detailed derivation of (2.1) but, rather, to put (2.1) in context with Fig. 2.14.

Some key insights on the behavior of phase noise with respect to OFDM can be obtained by zooming in on a specific symbol r_j whose expression is given by

$$r_j = \Big(\sum_{i=0}^{N_c-1} \delta^{\mathrm{R}}_{i-j} H_i \delta^{\mathrm{T}}_{-i+j} \Big) s_j + \sum_{k=0,k\neq j}^{N_c-1} \Big(\sum_{i=0}^{N_c-1} \delta^{\mathrm{R}}_{i-j} H_i \delta^{\mathrm{T}}_{-i+k} \Big) s_k + w_j. \qquad (2.4)$$

From (2.4), we see that the desired symbol s_j is corrupted by two terms: A multiplicative distortion term given by $\Big(\sum_{i=0}^{N_c-1} \delta^{\mathrm{R}}_{i-j} H_i \delta^{\mathrm{T}}_{-i+j} \Big)$, also known as *common phase error* (CPE); The second term is known as the *inter-carrier interference* (ICI) which is an added noise contribution to s_j and is given by second term in (2.4). In order to appreciate the effect of phase noise, we need to see how (2.4) reduces when there is no phase noise. In the absence of the phase noise, we have that $\mathbf{V}_{\mathrm{x}} = \mathbf{I}_{N_{\mathrm{c}}}$, where $\mathbf{I}_{N_{\mathrm{c}}}$ is the $N_{\mathrm{c}} \times N_{\mathrm{c}}$ identity matrix. Equation (2.1), thus, reduces to $\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{w}$ and, hence, $r_j = H_j s_j + w_j$; That is there is no more any CPE or ICI.

2.4 Phase Noise Modeling

Modeling physical phenomena and systems that process them provides valuable insight and understanding about their behavior. In this section, we concern ourselves with modeling the phase noise process. Phase noise modeling (more precisely, we model the oscillators) helps establishes the relation between certain key oscillator parameters and the phase noise process itself. Typically, the oscillator device is made up of noisy integrated-circuit components, and the goal in modeling is to understand or, at the least, provide a relation between these various noise sources and the resulting phase noise process.

Oscillators are physical and, in most cases, man-made devices that produce *periodic signals*. The periodic nature of the oscillator signal provides for excellent *time-keeping* which is essential in digital systems. With regard to telecommunications and in addition to providing a time reference, they are also used for *information transmission*, i.e., the oscillator signal acts as an information carrier. In this thesis, we concern ourselves with only sinusoidal signals generated by an oscillator device. An ideal oscilla-

tor produces signals of the form

$$o(t) = e^{j2\pi f_c t}, \tag{2.5}$$

where f_c denotes the frequency of the exponential sinusoid. In the context of telecommunications, f_c would be the carrier frequency. In (2.5), we use, for convenience, the complex-representation of a sinusoid rather than the real-representation which is what oscillator devices physically produce. Unfortunately, oscillators are physical devices and, thereby, are prone to inherent *noise sources* that exist in the components that make up the oscillator. These noise sources eventually render a noisy version of $o(t)$ which mathematically is expressed as

$$o(t) = (1 + a(t))e^{j(2\pi f_c t + \theta(t))}, \tag{2.6}$$

where $\theta(t)$ denotes phase noise and $a(t)$ denotes the *amplitude noise*. In the circuit design community, these also go by the names of PM and AM referring to the phase and amplitude modulation of the carrier, respectively. It is shown in [54] that if the various noise sources in the oscillator are small then the amplitude noise is also bounded and small. With this fact in mind, in the rest of this section and thesis, we ignore the effect of $a(t)$. Of course, plenty of studies are available that seek to characterize amplitude noise. We refer the interested reader to [19, 55] and [56].

2.4.1 The Power-Law Model

By modeling of phase noise, we imply either a time domain characterization of the signal $\theta(t)$ or a spectral characterization, for example, through the power spectral density (PSD). Denote the respective PSD of $\theta(t)$ and $o(t)$ by $S_\theta(f)$ and $S_o(f)$. In the open literature, the popular *power-law* model for $S_\theta(f)$ is used which generically is given by

$$S_\theta(f) = \sum_{i=0}^{4} \frac{c_i}{f^i}, \tag{2.7}$$

where constants c_i are generally obtained numerically or using actual measurement data. The justification for the use of (2.7) is mainly driven from experimental data on $S_o(f)$, where various power-law terms dominate certain regions of the spectrum [57, 58]. The relation between $S_o(f)$ and $S_\theta(f)$ is given by [59]

$$S_\theta(f) \approx L(f) = \frac{S_o(f_c + f)}{P_s},\ f \gg 0, \tag{2.8}$$

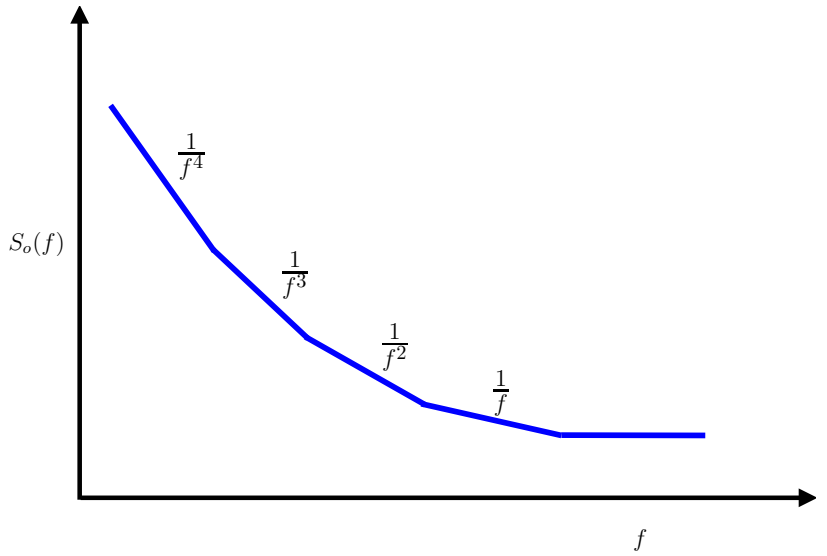


Figure 2.15. Illustration of the PSD of a practical oscillator.

where $L(f)$ is known as the single-side-band (SSB) phase noise spectrum, and $P_{\rm s}$ is the carrier signal power. From (2.8), the equivalence between $S_o(f)$ and $S_\theta(f)$ is valid only for large frequency offsets from the carrier frequency. In general, the quantity $L(f)$ is obtained experimentally and, after using (2.8), the constants c_i of the power-law model in (2.7) can be obtained. In Fig. 2.15, an illustration of the power-law model is shown. The PSD is typically characterized by a noise floor (white noise process) for large frequency offsets from the carrier frequency.

The power-law model, in general, predicts well the shape of the spectral density of $S_o(f)$ for large frequency offsets. However, as seen from (2.7), it blows up to infinity for near-carrier frequencies which is impossible as the oscillator process is always stationary and, thus, has finite-power [54]. A hybrid model that combines a different model for near-carrier frequencies and the power-law model for large frequency offsets is indeed the way forward to a more accurate description of the phase noise spectral characteristic. Some recent works in these areas can be found in [58, 60, 61] and references therein. Nevertheless, the power-law model is extremely useful for predicting phase noise behavior. In the following sections, we discuss some general findings on the relation between this power-law characteristic and the various noise sources that exist in the oscillator device.

2.4.2 Phase Noise in Free-Running Oscillators

The term *free-running* refers to the fact that in such oscillators the phase is allowed to drift, i.e., $\theta(t)$ grows unbounded. In most practical situations, such oscillators are never used in isolation rather they are used in a feedback loop by means of a PLL which is the subject of the next section.

Free-running oscillators are typically characterized by various sources of noise that eventually get modulated in a non-linear fashion to render the undesired phase noise [57, 62]. Some of these noise types are, for example, shot noise, thermal noise, burst noise and the so-called $1/f$ noise. Shot and thermal noise are best modeled as white noise processes, while burst and $1/f$ noise have a colored density spectrum [63]. In [54] and [63], the effect of these noise sources on $\theta(t)$ and the resulting oscillator PSD is rigorously analyzed. Using nonlinear perturbation analysis and *Floquet theory*, it is shown that, asymptotically, $\theta(t)$ is a zero-mean Gaussian process with variance that generally grows with time. Specifically, as shown in [54], for white noise sources, $\theta(t)$ is a *Wiener process* whose variance is given by

$$\sigma_\theta^2(t) = c_\mathrm{w} t, \tag{2.9}$$

where the constant c_w characterizes the white noise source. The oscillator PSD for Wiener phase noise follows a *Lorentzian* and takes the form [54]

$$S_o(f_\mathrm{off}) \propto \frac{f_\mathrm{c}^2 c_\mathrm{w}}{f_\mathrm{c}^4 c_\mathrm{w}^2 + (f_\mathrm{off})^2}, \tag{2.10}$$

where f_off represents the offset from f_c. As can been seen from (2.10), the PSD has finite power at $f_\mathrm{off} = 0$ and, in fact, is nearly flat for frequencies close to f_c.

In [63], the above results are extended to also include colored noise sources. Specifically, the oscillator PSD is given by

$$S_o(f_\mathrm{off}) \propto \begin{cases} \frac{f_c^2 (c_\mathrm{w} + S_\mathrm{N}(0))}{f_\mathrm{c}^4 (c_\mathrm{w} + S_\mathrm{N}(0))^2 + (f_\mathrm{off})^2}, & f_\Delta \approx 0 \\ \frac{f_\mathrm{c}^2}{f_\mathrm{off}^2} \left(c_\mathrm{w} + S_\mathrm{N}(f_\mathrm{off}) \right), & f_\mathrm{off} \gg 0 \end{cases} \tag{2.11}$$

In the above equation $S_\mathrm{N}(f_\mathrm{off})$ denotes the spectral density of the colored-noise source. Typical colored noise models assume the power-law characteristic for $S_\mathrm{N}(f_\mathrm{off})$ and, after using (2.11), for large frequency offsets, other power-law factors can be obtained. In that sense, (2.11) validates the use of the empirical power-law model. Using similar methods adopted in [63], the work in [56] includes the effect of amplitude noise when determining the oscillator PSD.

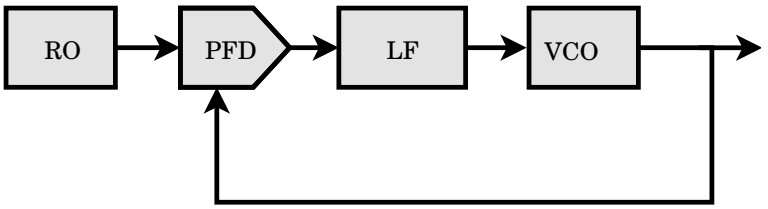


Figure 2.16. Basic components constituting a PLL.

2.4.3 Phase Noise in Phase-Locked Loop Based Devices

The (unbounded) phase drift encountered in oscillators may not be acceptable in many digital applications. To overcome this problem, some form of feedback mechanism is always employed to keep the oscillator phase bounded. In that respect, PLLs are widely used, where the oscillator in question is *phase locked* to a reference higher quality oscillator. A typical PLL schematic is shown in Fig. 2.16. The phase of a voltage controlled oscillator (VCO) is compared with a high-quality reference oscillator (RO) by means of a phase-frequency detector (PFD). The phase difference is then low-pass filtered by a loop filter (LF) whose output drives the VCO. The PLL is said to be *locked* when the RO frequency and VCO frequency are equal and the phase difference is constant.

The sources of noise that creep into the PLL-device are the phase noise from the free-running RO and VCO, noise sources from the PFD and the loop filters. All these sources of noise and the construction of the PLL device affect the resulting phase noise seen at the output of the VCO. In [64], building upon the work of [54] and using non-linear perturbation analysis, the work aims to obtain a time-domain characterization of the phase noise process in a PLL-device and also determine the PSD of the PLL output. By assuming *white* noise sources and the PLL being in locked condition, the work in [64] demonstrates that the resulting phase noise in a PLL-device is a Gaussian process and can be expressed as a sum of the RO Wiener process and one component of a multi-dimensional *Ornstein–Uhlenbeck process* [65]. Closed-form analytical expressions of the oscillator PSD are also derived which can be found in [64]. Although not obvious from the derived expressions of the PSD, through examples, it is seen that, for low frequency offsets from the carrier frequency, the PSD is equal to the PSD of the free-running RO, while for large offsets, it is equal to the PSD of the free-running VCO.

LTI Phase-Domain Model

The mathematical characterization of the PLL used in [64] is based on the ordinary differential equation (ODE) formulation of the PLL-mechanism. It is an accurate description of the PLL that takes into account the inherent non-linearities that shape the eventual phase noise process. However, the resulting characterization requires numerical techniques to compute the various parameters of the model which can be quite cumbersome.

In the general literature related to phase noise analysis for PLL devices, a linear time-invariant (LTI) approach is used, where the various noise sources propagate in the *phase domain*. Figure 2.17 shows the phase-domain LTI model for a *charge-pump* (CP) PLL [66]. A CP-PLL uses a charge-pump device between the PFD and the LF. It delivers a current charge rather than voltage to the LF. CP-PLLs are widely used in digital systems and are a popular choice among various PLL types available. They are well known for their flexible design allowing to trade between different design parameters. We refer the reader to [67,68] and references therein for a better and more detailed treatment.

The second part of Fig. 2.17 shows the corresponding LTI phase-domain model. The model includes noise sources from the PFD, LF, RO and VCO. Also represented in the figure is the noise from the frequency dividers. The loop filter, LF, is typically a low pass filter whose bandwidth can be controlled by varying the LF resistance R_s and the capacitance C_1 and C_2. The LTI system representing the CP device is simply the gain factor of $I_\mathrm{p}/2\pi$, where I_p represents the current delivered to the LF. In the second part of Fig. 2.17, an ideal integrator with transfer function $\frac{K_\mathrm{VCO}}{s}$ follows the LF. Such a block is used in the model by recognizing that the frequency of a VCO is controlled by its input voltage and, thus, its phase is obtained by performing an integration operation.

Using this LTI phase-domain model, the corresponding transfer functions seen by the various noise sources can be easily derived and is given by [66]

$$H_\mathrm{VCO}^\mathrm{PLL}(f) = \frac{\Theta_\mathrm{VCO}^\mathrm{PLL}(f)}{\Theta_\mathrm{VCO}(f)} = \frac{\jmath 2\pi f N}{g(f)}, \tag{2.12}$$

$$H_\mathrm{RO}^\mathrm{PLL}(f) = \frac{\Theta_\mathrm{RO}^\mathrm{PLL}(f)}{\Theta_\mathrm{RO}(f)} = \frac{N I_p K_\mathrm{VCO} H_\mathrm{LF}(f)}{M g(f)}, \tag{2.13}$$

$$H_\mathrm{PFD}^\mathrm{PLL}(f) = \frac{\Theta_\mathrm{PFD}^\mathrm{PLL}(f)}{\Theta_\mathrm{PFD}(f)} = \frac{2\pi N K_\mathrm{VCO} H_\mathrm{LF}(f)}{g(f)}, \tag{2.14}$$

$$H_\mathrm{LF}(f) = \frac{1 + \jmath 2\pi f R_s C_1}{\jmath 2\pi f (C_1 + C_2 + \jmath 2\pi f R_s C_1 C_2)}, \tag{2.15}$$

where $\Theta_\mathrm{W}^\mathrm{PLL}(f)$, $\mathrm{W} \in \{\mathrm{VCO}, \mathrm{RO}, \mathrm{PFD}\}$, denotes the frequency response

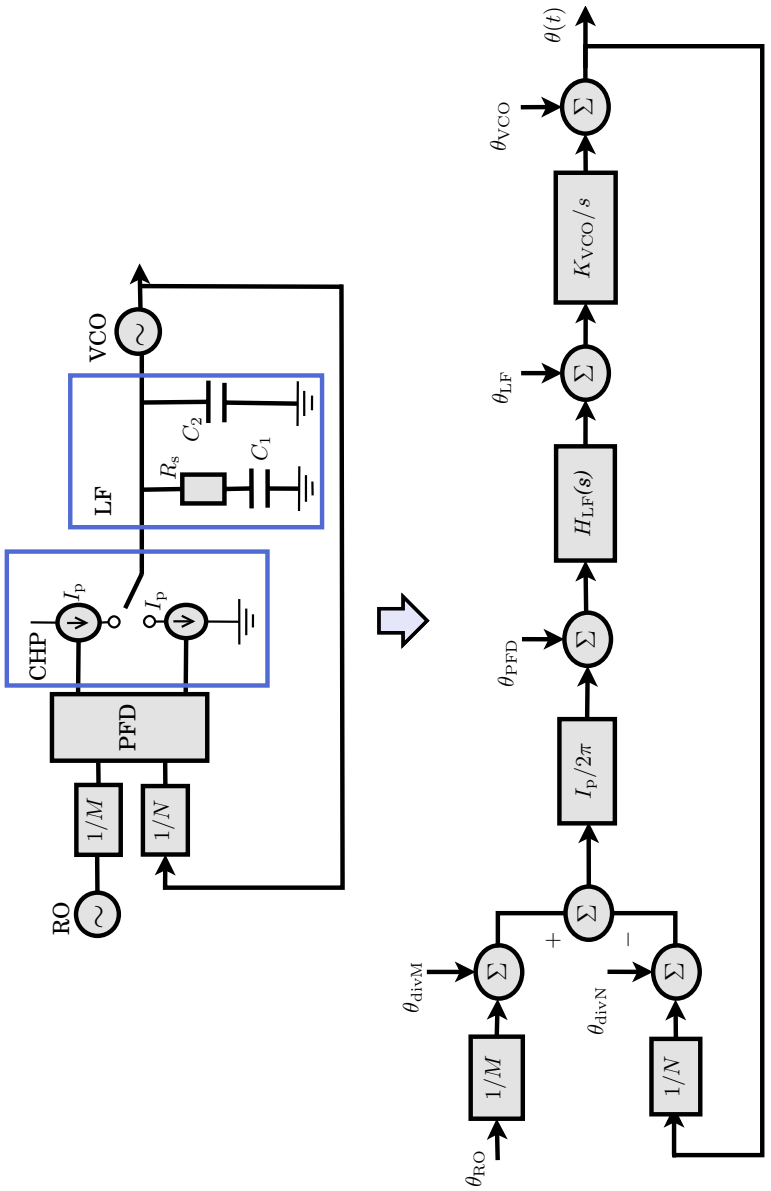


Figure 2.17. Charge-pump PLL and the corresponding LTI phase-domain model.

of the output, $\theta_{\mathrm{W}}^{\mathrm{PLL}}(t)$, corresponding to the input $\theta_{\mathrm{W}}(t)$ whose frequency response is denoted by $\Theta_{\mathrm{W}}(f)$. The expression for $g(f)$ is given by

$$g(f) = \jmath 2\pi f N + I_{\mathrm{p}} K_{\mathrm{VCO}} H_{\mathrm{LF}}(f). \tag{2.16}$$

In deriving the above equations, we assume that the noise sources are only in the VCO, RO and PFD. By applying bi-linear transformation, we can transform the continuous-time LTI system to a parallel discrete system as shown in Fig. 2.18. In this thesis, we refer to such a system as

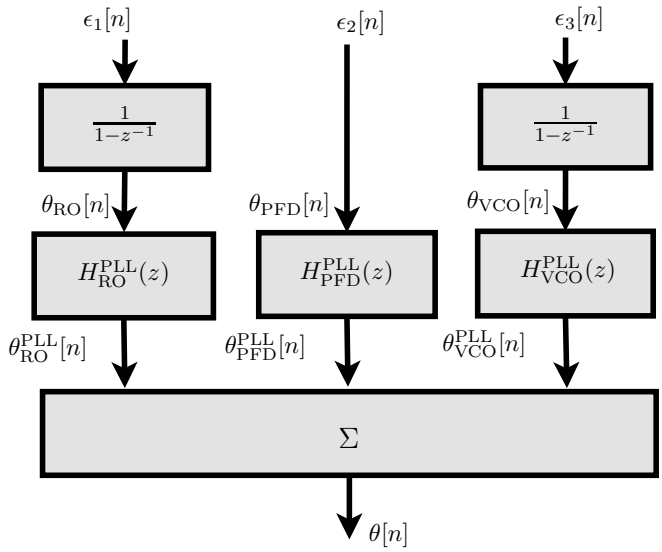


Figure 2.18. The discrete PARMA model.

a discrete parallel auto regressive moving average model (PARMA). The inputs $\theta_{\mathrm{W}}[n]$ are a discretized version of their continuous-time counterparts. These discretized inputs are obtained by passing white Gaussian noise through suitable input filters. For example, in Fig. 2.18, we assume that $\theta_{\mathrm{PFD}}[n]$ is white noise, while $\theta_{\mathrm{RO}}[n]$ and $\theta_{\mathrm{VCO}}[n]$ are discrete Wiener processes which can be obtained by passing white Gaussian noise through an LTI system with transfer function $H(z) = \frac{1}{1-z^{-1}}$.

In Fig. 2.18, we have only considered three noise sources which correspond to three parallel branches in the discrete PARMA model of PLL phase noise. In general, we can extend the analysis to include other noise sources which correspond to adding more parallel branches. Assuming L parallel branches and denoting the overall impulse response of the pth branch by $h_p[n]$, we have its corresponding transfer function given by

$$H_p(z) = \frac{a_{p_0} + a_{p_1} z^{-1} + a_{p_2} z^{-2} + \ldots + a_{p_V} z^{-V}}{1 + b_{p_1} z^{-1} + b_{p_2} z^{-2} + \ldots + b_{p_V} z^{-V}}. \tag{2.17}$$

With the above transfer function, the PLL phase noise is obtained as

$$\theta[n] = \sum_{p=1}^{L} \theta_p[n] = \sum_{p=1}^{L} \sum_{j=1}^{\infty} \epsilon_p[n-j+1] h_p[j], \tag{2.18}$$

where $\theta_p[n]$ denotes the output of pth parallel filter. The L independent zero-mean and unit variance white Gaussian inputs to the filters are denoted by $\{\epsilon_p[n]\}_{p=1}^{L}$. The term $h_p[j]$ can be obtained recursively as follows: $h_p[1] = a_{p_0}$; and $h_p[j] = a_{p_{(j-1)}} - \sum_{k=1}^{V} b_{p_k} h_p[j-k]$.

2.5 Summary

This chapter presented a brief motivation of the problem that this thesis seeks to address by setting the appropriate background. In particular, the direct-conversion architecture for the analog back-end unit was presented. It is an architecture that performs frequency conversion directly between RF and baseband frequency. It is a popular choice for the analog back-end unit mainly for its system on-chip integration capability which is advantageous for its low power consumption and smaller form factor. The major functionalities associated with this type of architecture are power amplification, frequency conversion and analog-to-digital/digital-to-analog conversion, where these functionalities are implemented by the respective power amplifiers, oscillators and analog-to-digital and digital-to-analog converters. Unfortunately, these electronic devices, like any physical system, have their limitations and are prone to noise and, thus, the functionalities are far from the idealistic scenario. The realistic scenario is the distortion of the transmitted and received signal.

This thesis considers only the problem encountered during frequency conversion in the analog back-end unit. Specifically, the problem of phase noise is considered. It is assumed in this thesis that problems related to power amplification and analog-to-digital/digital-to-analog conversion are non-existent or somehow taken care of. At this point, the reader is informed that there is plethora of research that addresses these other problems of the analog unit. We refer the reader to see the references pertaining to the relevant sections of this chapter.

The second part of this chapter is primarily about modeling. A linear model of the communication link under the influence of both transmit and receive phase noise is presented. The communication link employs the very effective OFDM modulation scheme which essentially is a technique of packing information with certain objectives in mind. The basic principle of OFDM is discussed in this chapter, and its use in practice is also presented. This linear model of the communication link serves as the foundation to conduct performance analysis and at the same time in developing phase noise estimation algorithms. Finally, the last part of this chapter is on phase noise modeling. A brief treatment of the commonly used phase noise models, especially, those that occur in free-running and PLL-based oscillators are discussed. These models are standard and widely used in the research community in general.

3. Characterization and Analysis of OFDM with Phase Noise

Phase noise destroys the orthogonality between subcarriers of an OFDM signal thereby resulting in each subcarrier experiencing interference from its neighboring subcarriers. One can, thus, expect performance degradation compared to the case of a phase noise-free OFDM link. Extensive work, in the past two decades, has been devoted towards quantifying this performance degradation which is the subject of this chapter. In the first part of this chapter, we review some of the prior work on performance analysis for OFDM systems impaired by phase noise, carrier frequency offset or both. Typical measures of performance are the signal-to-interference-plus-noise ratio (SINR), symbol-error probability (SEP), bit-error probability (BEP) and channel capacity. All of these works, overwhelmingly, demonstrate a performance drop for an OFDM system impaired by either phase noise or frequency offset or both.

The second part of this chapter is devoted towards the contributions of this thesis in relation to the topics of characterization and analysis. Specifically, the results of Publications I–IV are briefly reviewed. The works in Publications I, II and III have a common underlying goal which is to assess the degradation in channel capacity. In determining the capacity, interesting characterizations, related to the phase noise process, are also discovered. For example, for slow-varying phase noise processes, the PDF of the so-called *ICI power* can be approximated as a sum of correlated gamma random variables. Phase noise affects the channel capacity through this ICI power. The work in Publication IV is mainly about characterization, where PDF of the spectral components of the complex exponential of phase noise are the subject of study. It is shown that, for slow-varying phase noise processes, the spectral components can effectively be represented as the sum of independent Gaussian and gamma-like distributed random variables.

3.1 Prior Work on Phase Noise Analysis for OFDM

The performance of communication systems are typically evaluated using metrics such as signal-to-noise ratios (SNR), BEP, SEP and channel capacity [69]. Some of the earliest studies investigating the performance of phase noise impaired OFDM demonstrated performance drops in these metrics [70–73]. These studies, however, were based on numerical simulations of the performance metrics and, hence, do not necessarily provide qualitative insight on the relationship between these metrics and the undesired phase noise and frequency offset. Effort was, thus, directed towards determining closed-form analytical expressions of these performance metrics.

The works were initially focused on determining the degradation in SINR and gradually progressed towards determining BEP and SEP. The general expression for the SINR can be derived from the OFDM system model impaired by both transmit and receive phase noise of (2.4). Restated here, the system model is given by

$$r_j = \Big(\sum_{i=0}^{N_c-1} \delta^R_{i-j} H_i \delta^T_{-i+j} \Big) s_j + \sum_{k=0,k\neq j}^{N_c-1} \Big(\sum_{i=0}^{N_c-1} \delta^R_{i-j} H_i \delta^T_{-i+k} \Big) s_k + w_j. \tag{3.1}$$

In the above equation, the desired symbol s_j is corrupted by the multiplicative distortion term, also known as CPE, and an extra additive noise term, also known as ICI, which represents the interference from other symbols s_k. By evaluating the average powers of the above terms, the SINR is given by

$$\bar{\bar{\Upsilon}}_j = \frac{P_{\text{cpe}} \sigma_{\text{s}}^2}{P_{\text{ici}} \sigma_{\text{s}}^2 + \sigma_{\text{w}}^2}, \tag{3.2}$$

where the average CPE power and ICI power is given by

$$P_{\text{cpe}} = \mathrm{E}\left\{ \Big| \sum_{i=0}^{N_c-1} \delta^R_{i-j} \mathcal{H}_i \delta^T_{-i+j} \Big|^2 \right\}, \tag{3.3}$$

$$P_{\text{ici}} = \sum_{k=0,k\neq j}^{N_c-1} \mathrm{E}\left\{ \Big| \sum_{i=0}^{N_c-1} \delta^R_{i-j} \mathcal{H}_i \delta^T_{-i+k} \Big|^2 \right\} \tag{3.4}$$

with $\mathrm{E}\{\cdot\}$ denoting the expectation operator. The respective signal and noise powers are $\sigma_{\text{s}}^2 = \mathrm{E}\{|s_j|^2\}$ and $\sigma_{\text{w}}^2 = \mathrm{E}\{|w_j|^2\}$. In the absence of phase noise, we have $P_{\text{cpe}} = 1$ and $P_{\text{ici}} = 0$, and (3.2) reduces to

$$\bar{\bar{\Upsilon}}_j = \frac{\mathrm{E}\{|H_j|^2\} \sigma_{\text{s}}^2}{\sigma_{\text{w}}^2} \tag{3.5}$$

which is the average SNR per subcarrier without phase noise.

Much effort in the general literature has been devoted towards evaluating meaningful expressions for P_{cpe} and P_{ici}. One of the earliest works in doing so was in [74], where the authors derive closed-form expressions assuming an additive white Gaussian noise (AWGN) channel, i.e., $\{H_k\}_{k=0}^{N_c-1} = 1$ while also including the effect of carrier frequency offset. This work was extended to multipath channels in [75], however, the multipath channel was a two-tap channel impulse response with equal amplitude. In [76], a second-order approximation of the complex exponential of the phase noise is used in arriving at approximations for P_{cpe} and P_{ici}. A more accurate representation of P_{cpe} and P_{ici}, for Wiener phase noise, is derived in [77]. Equation (3.2) can be interpreted as the SINR without CPE compensation. In [78], the authors extend the SINR analysis to receivers that perform CPE compensation and also receivers with differential signaling.

The average SINR in (3.2) can be used to obtain the metrics of BEP and SEP (equivalently bit-error rate (BER) and symbol-error rate (SER)). Such an approach, for example, has been utilized in the works of [74, 76, 77]. However, this may not necessarily yield the exact BEP or SEP. In general, for any given signal constellation and in the absence of phase noise, the BEP and SEP depend upon the set SNR, where the receiver noise is assumed to be white Gaussian [69]. By using the SINR expression of (3.2) in these Gaussian-based expressions of BEP, approximate BEP and SEP can be obtained for a phase noise impaired OFDM system. Another approach is to assume that the ICI term in (3.1) is Gaussian and derive, from first principles, the BEP and SEP for various signal constellations as done in [79] and [80].

The Gaussian assumption for the ICI term no doubt renders tractable and neat mathematical expressions of BEP and SEP. The assumption, however, is questioned in [78], where the authors investigated the cases where one can use the SINR of (3.2) in the Gaussian-based BEP or SEP expressions. It was speculated that for slow-varying phase noise processes the ICI term is not Gaussian while, for fast-varying ones, it is indeed Gaussian. In [81], a semi-analytical approach is used to determine the SEP for BPSK, QPSK and 16-QAM constellations without any Gaussian assumption for the ICI term. The work is based on using the *Beaulieu series*-based expansion for SEP derived for an OFDM system impaired by frequency offset of [82]. In what can be regarded as seminal work, the authors in [52, 83] investigate the Gaussian hypothesis for the ICI term in

(3.1), however, under the assumption of an AWGN channel. The authors derive the asymptotic PDF of the ICI term and show that, for slow-varying phase noise processes, it exhibits much thicker tails than the Gaussian-based PDF assumption.

The review of prior work, performed in this section, is mainly focused towards phase noise analysis for OFDM. By no means is this review complete. In fact, numerous works are available that study the joint effects (or some combination of it) of various RF-impairments in general. These include phase noise, jitter noise, frequency offset, IQ-imbalance and power amplifier non-linearities. The reader may refer to some works of [38, 39, 84–91] and references therein.

3.2 Contributions to Phase Noise Analysis for OFDM

This thesis presents new results derived in Publications I–IV on characterization and closed-form analytical expressions of channel capacity for OFDM systems impaired by phase noise and carrier frequency offset in multi-path fading channels. Starting with Publication I, the channel capacity is derived assuming a Wiener phase noise model, while in Publication II, the analysis is extended to PPL-based phase noise processes which is modeled using a discrete PARMA model. These results are extended in Publication III to also include the effect of carrier frequency offset.

In addition to the final objective of determining the channel capacity, some new characterization of the phase noise variables are also discovered. For example, in Publications I and II, it is shown that, for slow-varying phase noise processes, the PDF of the so-called ICI power is a sum of correlated gamma random variables. In earlier literature, only second-order statistics of this ICI power were available. Results on its distribution when including carrier frequency offset are derived in Publication III. The derived ICI power PDF can also be used, for example, to obtain the capacity of the OFDM system impaired by both phase noise and IQ-imbalance as done in [92]. In Publication IV, for slow-varying phase noise processes, the PDF of the real and imaginary parts of the spectral components of the complex exponential of phase noise are derived. It is shown that they can effectively be represented as a sum of two independent random variables: The first is a stronger Gaussian random variable; and the second is a weaker gamma-like random variable.

3.3 Characterization and Analysis

We now summarize the works presented in Publications I–IV. Specifically, this summary is about the common underlying methodology adopted in these works. The specifics related to each of these works are mentioned at the appropriate places.

3.3.1 The Instantaneous SINR

Equation (3.2) is an estimate of the average SINR, where the expectations of the numerator terms and denominator terms are desired. As can be seen from (3.2), the numerator and denominator involve the same phase noise and channel variables. Thus, strictly speaking, (3.2) is not an accurate physical description of the average SINR. In order to obtain a more rigorous formulation, a *conditional* SINR must first be evaluated, where the channel and phase noise variables are conditioned on a fixed realization. This conditional SINR is given by

$$\Upsilon_j = \frac{\left|\sum_{i=0}^{N_c-1} \delta_{i-j}^R \mathcal{H}_i \delta_{-i+j}^T\right|^2 \sigma_{\mathrm{s}}^2}{\sum_{k=0,k\neq j}^{N_c-1} \left|\sum_{i=0}^{N_c-1} \delta_{i-j}^R \mathcal{H}_i \delta_{-i+k}^T\right|^2 \sigma_{\mathrm{s}}^2 + \sigma_{\mathrm{w}}^2}. \tag{3.6}$$

From (3.6), we see that the SINR, for any jth subcarrier, depends on a particular realization of the channel through $\mathcal{H}_i$ and the phase noise through δ_i^X. For different realizations of the channel and phase noise, we get different realizations of Υ_j and, thus, we see that the SINR can be described as a random variable.

An accurate approximation to (3.6) can be found based on the following assumptions: First, the oscillator PSD 3-dB bandwidth, denoted by $f_{3\mathrm{dB}}$, is much smaller than the subcarrier spacing which is denoted by f_{sub}; and second, the channel coherence bandwidth is much larger than f_{sub}. For example, oscillator PSD 3-dB bandwidth are in the range of a few hundreds of Hertz which is much smaller than the 15 kHz subcarrier spacing specified for LTE [47,93,94]. On the other hand, the coherence bandwidth of wireless channels are in the order of several hundreds of kilo Hertz. In Publications I and II, these assumptions are used to arrive at a simpler expression for (3.6) and is given by

$$\Upsilon_j \approx \frac{|\delta_0|^2|\mathcal{H}_j|^2\sigma_{\mathrm{s}}^2}{|\mathcal{H}_j|^2\Big(\sum_{k=1}^{N_c-1}|\delta_k|^2\Big)\sigma_{\mathrm{s}}^2 + \sigma_{\mathrm{w}}^2} = \frac{1-Y}{Y + \frac{\sigma_{\mathrm{w}}^2}{\sigma_{\mathrm{s}}^2 G_j}}. \tag{3.7}$$

The random variables Y and G_j characterize the phase noise and channel

respectively and are given as follows:

$$Y = \sum_{k=1}^{N_c-1} |\delta_k|^2;\ G_j = |\mathcal{H}_j|^2, \tag{3.8}$$

where the coefficient δ_k denotes the DFT of the combined transmit and receiver phase noise, i.e.,

$$\delta_k = \frac{1}{N_c} \sum_{n=0}^{N_c-1} e^{\jmath(\theta^T[n]+\theta^R[n])} e^{-\jmath(2\pi kn)/N_c}. \tag{3.9}$$

Equation (3.7) is useful compared to (3.6) in the following ways: It is simpler as it is characterized by only two independent random variables Y and G_j; It provides some insight on the relation between Υ_j and phase noise. In the absence of phase noise, we have $\delta_0 = 1$ and $\delta_k = 0$ for $k > 0$ which implies $Y = 0$ and, hence,

$$\Upsilon_j = \frac{|\mathcal{H}_j|^2 \sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{w}}^2} \tag{3.10}$$

which is the SNR of an OFDM radio link without phase noise. However, in the presence of phase noise, we have $Y > 0$ which results in a reduction from the phase noise-free case. In this thesis, we refer to the random variable Y as the 'ICI power'.

The SINR in (3.7) depends on particular realizations of the independent random variables Y and G_j. Thus, Υ_j in (3.7) is also a random variable, and any statistical measure based on it will require knowledge of the distributions of Y and G_j. The random variable G_j characterizes the channel, and its distribution is well defined assuming a complex Gaussian channel. The distribution of Y, however, is not obvious at cursory glance and, in order to determine the distribution, the characterization of Y in terms of known well-defined elements is required. In Publications I and II, such a characterization is sort after using which the distribution of Y is derived.

3.3.2 Characterization of ICI Power

The characterization of the ICI power is first investigated in Publication I for the Wiener phase noise process, and in Publication II, the result is extended to the discrete PARMA phase noise model for PLL-based phase noise processes. Using Taylor series approximation, the ICI power is expressed as a sum of correlated gamma random variables and is given by

$$Y \approx \sum_{l=1}^{N_c-1} \sum_{i=1}^{N_c-l} Z_{il}, \tag{3.11}$$

where

$$Z_{il} = \frac{1}{N_c^2}\Big(\Delta\theta[i,l]\Big)^2, \tag{3.12}$$

$$\Delta\theta[i,l] = \theta[i+l-1] - \theta[l-1]. \tag{3.13}$$

The combined transmit and receive phase noise process is

$$\theta[n] = \theta^T[n] + \theta^R[n]. \tag{3.14}$$

Typically, $\theta[n]$ is characterized as a zero-mean Gaussian process. For example, for a discrete Wiener phase noise process, we have

$$\theta[n] = \sum_{i=1}^{n} \varepsilon(i), \tag{3.15}$$

where the i.i.d. $\varepsilon(i)$ are zero-mean Gaussian with some variance. This implies that $\Delta\theta[i,l]$ is also zero-mean Gaussian distributed and, hence, Z_{il} follows a gamma distribution with parameters $\alpha = \frac{1}{2}$ and $\beta = \beta_{il}$ which is a function of the variance of $\Delta\theta[i,l]$. Thus, we see that Y can be characterized by the sum of correlated gamma random variables.

3.3.3 PDF of Sum of Gamma Variates

The gamma random variables in (3.11) have a nice structure in the sense that the α parameter for all of them is the same and is equal to $\frac{1}{2}$. This arises due to the assumption that $\theta[n]$ is zero-mean Gaussian. All the more, the random variables Z_{il} are correlated. This is because $\Delta\theta[i,l]$ is constructed from a set of N_c random variables $\{\theta[n]\}_{n=0}^{N_c-1}$ which in turn can be described using a finite set of independent Gaussian random variables. In Publication I, the PDF of a sum of correlated gamma random variables is derived using the *Moschopoulos technique* [95]. The PDF derived in Publication I is a generalization of the result of [96] which is applicable only for full-rank covariance matrix of the gamma variables. In (3.11), the gamma variables have a rank-deficient covariance matrix. The following theorem summarizes the result.

Theorem 3.3.1. *Let $\{Z_n\}_{n=1}^{N}$ be a set of N correlated gamma variates ($Z_n \sim \mathcal{G}(\alpha,\beta_n)$) with normalized covariance matrix $\mathbf{M_z}$ of any rank $R \leq N$. Then, the PDF of $Y = \sum_{n=1}^{N} Z_n$ is given as*

$$p_Y(y) = \prod_{n=1}^{R}\left(\frac{\lambda_1}{\lambda_n}\right)^{\alpha}\sum_{k=0}^{\infty}\left(\frac{\zeta_k y^{R\alpha+k-1}e^{\frac{-y}{\lambda_1}}}{\lambda_1^{R\alpha+k}\Gamma(R\alpha+k)}\right), \tag{3.16}$$

where $\{\lambda_n\}_{n=1}^{R}$ *are the ordered eigenvalues of the matrix* $\mathbf{PBP}^{\mathrm{T}}\mathbf{\Delta}$ *with* λ_1 *being the minimum. The* $\mathbf{P}$ *and* $\mathbf{\Delta}$ *matrices are obtained from eigenvalue decomposition of* $\mathbf{M_x}$ *which is related to* $\mathbf{M_z}$ *as*

$$(\mathbf{M_x})_{ij} = \sqrt{(\mathbf{M_z})_{ij}},\ i,j = 1,2,\ldots,N, \tag{3.17}$$

$$\mathbf{M_x} = \mathbf{C\Sigma C}^{\mathrm{T}},\ \ \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Delta}_{R\times R} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{3.18}$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \ldots & \mathbf{c}_R & \mathbf{\Omega}_1 & \mathbf{\Omega}_2 & \ldots & \mathbf{\Omega}_{N-R} \end{bmatrix}, \tag{3.19}$$

$$\mathbf{P} = [\mathbf{c}_1\ \mathbf{c}_2\ \ldots\ \mathbf{c}_R]^{\mathrm{T}},\ \ \mathbf{B} = \mathrm{diag}(\beta_1\ \beta_2\ \ldots\ \beta_N). \tag{3.20}$$

The weights $\zeta_k, k = 0,1,2,\ldots$*, are given as*

$$\zeta_0 = 1,\ \ \zeta_{k+1} = \frac{\alpha}{k+1}\sum_{i=1}^{k+1}\Big(\sum_{j=1}^{R}(1-\frac{\lambda_1}{\lambda_j})^i\Big)\zeta_{k+1-i}. \tag{3.21}$$

Proof. See Publication I. □

A careful observation of (3.16) provides a nice interpretation of the PDF of Y: Firstly, we note that the parenthesis term in (3.16) represents a gamma distributed PDF; and thus, the PDF of Y is expressed as a weighted sum of gamma distributed PDFs with weights ζ_k.

PDF of ICI Power under the Wiener Model

Theorem 3.3.1 can be used to determine the PDF of Y in (3.11). In Publication I, the parameters of the PDF of (3.16) for a Wiener phase noise model are derived. Of interest and importance is to relate the behavior of the PDF with the ratio $\rho = \frac{f_{3\mathrm{dB}}}{f_{\mathrm{sub}}}$ which is a measure of the level of inter-carrier interference. The behavior of the PDF in (3.16) is mainly dictated by the parameters R and λ_1 which is the smallest eigenvalue. In Publication I, for a Wiener phase noise model, it is shown that $R = N_{\mathrm{c}} - 1$ and

$$\lambda_1 \propto \frac{4\pi f_{3\mathrm{dB}} N_{\mathrm{c}}}{f_{\mathrm{sub}}} = 4\pi\rho N_{\mathrm{c}}, \tag{3.22}$$

where $\rho = \frac{f_{3\mathrm{dB}}}{f_{\mathrm{sub}}}$. From (3.22), we see that λ_1 increases linearly with ρ and, thus, we can expect a broadening in the PDF of Y. This is illustrated in Fig. 3.1, where we plot the PDF of Y for different values of N_{c} while keeping the system bandwidth and $f_{3\mathrm{dB}}$ fixed. Since, the bandwidth is kept constant, varying N_{c} implies varying f_{sub}, and hence, ρ also varies.

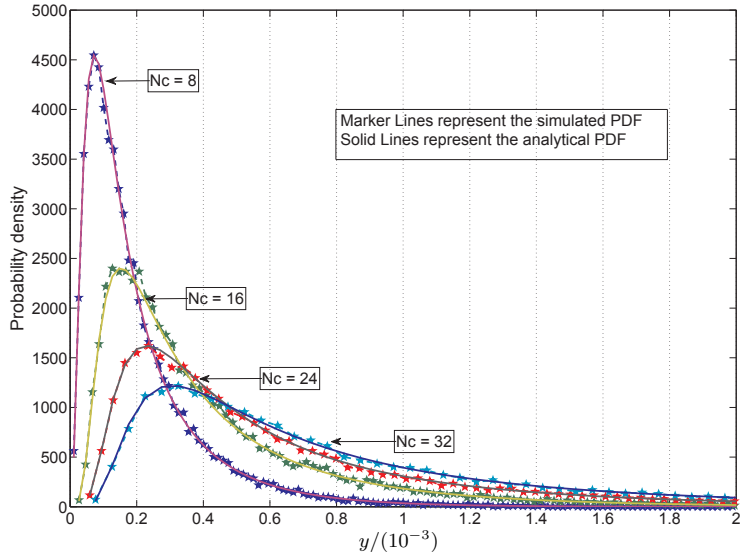


Figure 3.1. PDF of the ICI power for different values of N_{c}. The bandwidth of the OFDM system is set to 625 kHz and $f_{3\mathrm{dB}} = 200$ Hz.

PDF of ICI Power under the PARMA Model

For modeling phase noise in PLL-based devices, a discrete PARMA model is used which consists of a set of parallel auto-regressive-moving average filters. The resulting phase noise is given by

$$\theta[n] = \sum_{p=1}^{L} \theta_p[n] = \sum_{p=1}^{L} \sum_{j=1}^{\infty} \epsilon_p[n-j+1] h_p[j], \tag{3.23}$$

where $\theta_p[n]$ is the output from the pth parallel filter (with impulse response $h_p[n]$) corresponding to zero-mean, unit-variance white Gaussian inputs denoted by $\epsilon_p[n]$. We refer the reader to Chapter 2.4.3, where the relation between the filter coefficients and PLL device parameters is given. Utilizing this model, the expression for the gamma variables Z_{il} can be derived and is given by

$$Z_{il} \sim \mathcal{G}\left(1/2, \frac{2}{N_{\mathrm{c}}^2} \sum_{p=1}^{L} \sigma^2_{\eta_p^{il}}\right), \tag{3.24}$$

where the variance $\sigma^2_{\eta_p^{il}}$ is given by

$$\sigma^2_{\eta_p^{il}} = \sum_{j=1}^{i} h_p^2[j] + \sum_{j=1}^{N_p} \left(h_p[i+j] - h_p[j]\right)^2. \tag{3.25}$$

Since the diagonal matrix $\mathbf{B}$, defined in (3.20), is composed of elements

$$\beta_{il} = \frac{2}{N_{\mathrm{c}}^2} \sum_{p=1}^{L} \sigma^2_{\eta_p^{il}} \tag{3.26}$$

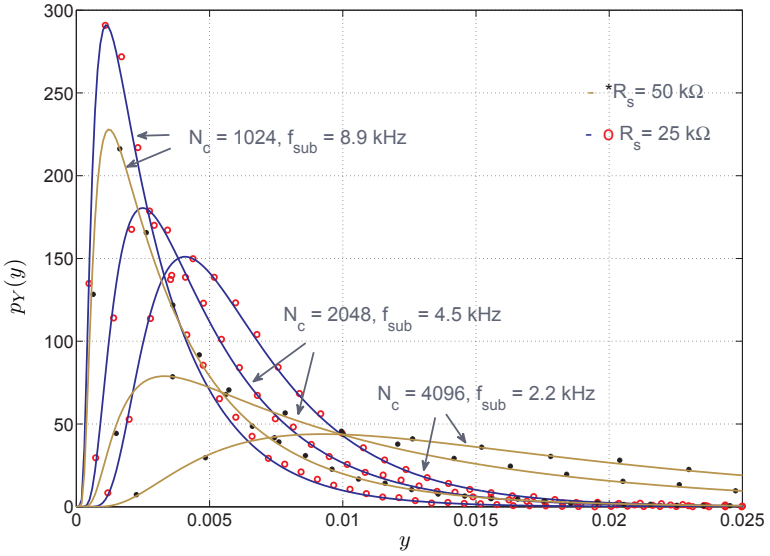


Figure 3.2. Comparison between simulated and analytical PDF plots of Y for different subcarrier spacings f_{sub} and phase noise PSD bandwidths obtained by varying the loop filter resistance R_{s}. The solid lines represent the analytical PDF while the marker lines represent the simulated histograms. The OFDM system bandwidth is 9.14 MHz. The PARMA filter parameters are given in Publication II.

and λ_1 is the eigenvalue of $\mathbf{PBP}^{\text{T}}\mathbf{\Delta}$, we have that

$$\lambda_1 \propto \sigma^2_{\eta_p^{il}}. \tag{3.27}$$

In Fig. 3.2, we plot the PDF of the ICI power for a PARMA phase noise model. The PARMA filter coefficients are obtained assuming a charge-pump PLL device. Such a device is shown in Fig. 2.17. The filter coefficients are obtained as described in Chapter 2.4.3. The phase noise bandwidth is controlled by varying the loop filter resistance R_{s} of the charge-pump PLL. An increase in R_{s} causes the PDF to spread over higher values of magnitude as seen in the figure. This behavior can also be explained using the PDF expression of (3.16). Firstly, the second term in (3.25) can be interpreted as the correlation between the impulse response coefficients of the pth parallel filter $h_p[j]$. Thus, for fast-varying phase noise processes (large values of R_{s}), we can expect less correlation between the coefficients of $h_p[j]$ and, thus, the second term is large which essentially results in a large value for λ_1 in (3.27). This effectively renders the PDF of Y towards higher values of magnitude.

Second-Order Statistics of ICI Power: The mean and variance of the ICI power can derived analytically using the PDF of (3.16). The mean

is evaluated as follows:

$$\begin{aligned}\bar{Y} = \mathrm{E}\{Y\} &= \int_0^\infty y p_Y(y) dy \\ &= K \sum_{k=0}^{\infty} \zeta_k \int_0^\infty \frac{y^{R\alpha+k} e^{\frac{-y}{\lambda_1}}}{\lambda_1^{R\alpha+k}\Gamma(R\alpha+k)} dy, \end{aligned} \tag{3.28}$$

where $K = \prod_{n=1}^{R} \left(\frac{\lambda_1}{\lambda_n}\right)^{\alpha}$. The integral above is of the form [97],

$$\int_0^\infty x^{v-1} e^{-\mu x} dx = \mu^{-v}\Gamma(v). \tag{3.29}$$

Applying (3.29) in (3.28), the final result for the mean is given by

$$\bar{Y} = K\lambda_1 \left(\sum_{k=0}^{\infty} \zeta_k \frac{\Gamma(R\alpha+k+1)}{\Gamma(R\alpha+k)} \right). \tag{3.30}$$

The variance of the ICI power is given by

$$\begin{aligned}\sigma_Y^2 &= \mathrm{E}\left[(y-\bar{Y})^2\right] = \mathrm{E}\left[y^2\right] - \bar{Y}^2, \\ &= \int_0^\infty y^2 p_Y(y) dy - \bar{Y}^2. \end{aligned} \tag{3.31}$$

Substituting the PDF of Y in (3.31) and making use of (3.29), we obtain

$$\sigma_Y^2 = K\lambda_1^2 \left(\sum_{k=0}^{\infty} \zeta_k \frac{\Gamma(R\alpha+k+2)}{\Gamma(R\alpha+k)} \right) - \bar{Y}^2. \tag{3.32}$$

From (3.30) and (3.32), we see that the mean and variance depends on the phase noise process through λ_1 which is in direct proportion to the level of phase noise. For fast-varying phase noise processes λ_1 takes higher values of magnitude which consequently imply larger values for the mean and variance of the ICI power.

3.3.4 Average Capacity

Channel capacity is a measure of the number of bits that can be transmitted through the channel with a very small error probability. In addition to SINR, SEP and BEP, channel capacity is a standard performance metric used in assessing the performance of a communication system. For an AWGN channel, the channel capacity is a function of the SNR or SINR in our case and is given by

$$\mathcal{C}_j = \log_2(1 + \Upsilon_j), \tag{3.33}$$

where Υ_j is the SINR seen by the jth subcarrier.

The capacity in (3.33) is applicable only under certain assumptions: *The total additive noise must be Gaussian*. In our case, the additive noise

given in (3.1), is the ICI plus the receiver Gaussian noise. As shown in [83], for slow-varying phase noise processes, the ICI is not Gaussian, thereby, the effective noise is non-Gaussian in general. Thus, in a strict information-theoretic sense, $\mathcal{C}_j$ is not the average capacity. However, it is the mutual-information assuming a *Gaussian input alphabet* for the transmitted symbols s_j. This is seen as follows: For a fixed realization of the channel and phase noise, the ICI plus receiver noise is Gaussian distributed. The SINR for this realization is given by Υ_j of (3.7) and, thus, the capacity for this Gaussian alphabet is obtained using (3.33). Such an approach of evaluating the channel capacity was also utilized in [98] for an OFDM link impaired by IQ-imbalance and in [92] for an OFDM link impaired by IQ-imbalance and phase noise. We shall, thus, refer to $\mathcal{C}_j$ of (3.33) as the capacity of (3.1) assuming a Gaussian input alphabet.

In Publications I and II, closed-form expressions of the average capacity are derived. It is obtained as follows: First $\mathcal{C}_j$ is averaged over the PDF of Y in (3.16), and the result is given by

$$\bar{\mathcal{C}}_j = \log_2\left(1+\gamma_j\right) - K\sum_{k=0}^{\infty}\zeta_k \log_2\left(1+b_k\gamma_j\right), \tag{3.34}$$

where $\gamma_j = \frac{g_j\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{w}}^2}$ denotes the instantaneous signal-to-noise ratio (SNR) with g_j being the realization of the random variable $G_j = |H_j|^2$ and

$$b_k = \frac{\Gamma(R/2+k+1)\lambda_1}{\Gamma(R/2+k)}. \tag{3.35}$$

The parameters K and ζ_k are defined in (3.16). Equation (3.34) is the capacity for a given realization of the channel g_j and has a nice interpretation: The first term represents the capacity without phase noise while the second term arises because of phase noise. Without any phase noise, the second term is zero since $\lambda_1 = 0$. However, in the presence of phase noise, we have a non-zero contribution from the second term, and the overall effect is a net-reduction from the phase noise-free case.

The average capacity, denoted by $\bar{\bar{\mathcal{C}}}_j$, is obtained by averaging $\bar{\mathcal{C}}_j$ over the PDF of g_j. Assuming a Rayleigh fading channel which implies that g_j is exponentially distributed, the final expression for the average capacity, derived in Publications I and II, is given by

$$\bar{\bar{\mathcal{C}}}_j = \log_2(e)\left[e^{\frac{1}{\bar{\gamma}_j}}E_1\Big(\frac{1}{\bar{\gamma}_j}\Big) - K\sum_{k=0}^{\infty}\zeta_k e^{\frac{1}{b_k\bar{\gamma}_j}}E_1\Big(\frac{1}{b_k\bar{\gamma}_j}\Big)\right], \tag{3.36}$$

where $E_1(\cdot)$ is the exponential integral function of order one and $\bar{\gamma}_j = \frac{\bar{g}_j\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{w}}^2}$ is the average SNR with $\bar{g}_j = \mathrm{E}\{g_j\}$. Similar to (3.34), the first term

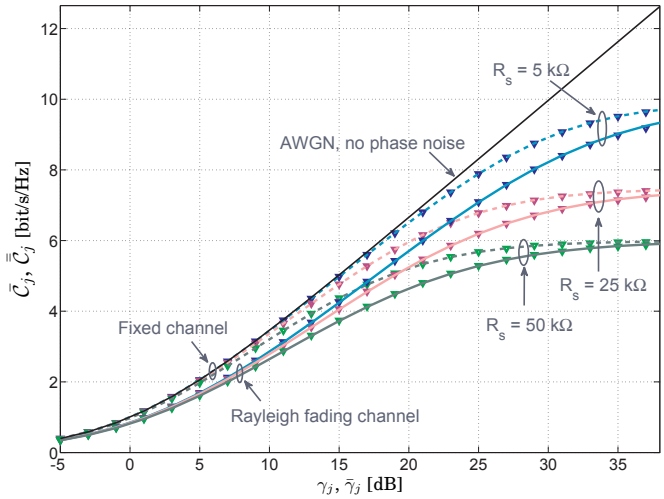


Figure 3.3. Capacity plots of $\bar{\mathcal{C}}_j$ and $\bar{\bar{\mathcal{C}}}_j$ for different phase noise PSD bandwidths obtained by varying the loop filter resistance R_s in the PARMA phase noise model. The markers and solid/dashed lines represent the simulations and analytical plots, respectively. The PARMA filter parameters are given in Publication II.

in (3.36) is the average capacity for a Rayleigh fading channel without phase noise, while due to phase noise, we have the non-zero second term in (3.36), and the result is a net-reduction from the Rayleigh fading capacity.

Figure 3.3 shows plots of the capacities of (3.34) and (3.36). The OFDM bandwidth is 9.14 MHz with $N_\mathrm{c} = 4096$. Channel is Rayleigh faded with 50 taps of exponential power delay profile with coherence bandwidth set to 400 kHz. The phase noise process used to generate these figures is of the discrete PARMA type. The phase noise bandwidth is controlled by varying the loop filter resistance R_s, and it increases with increase in R_s. As seen from Fig. 3.3, the presence of phase noise causes a reduction in the capacity when compared to the fixed channel case or Rayleigh fading case. This behavior of the capacity in (3.34) and (3.36) w.r.t. phase noise is through the parameter λ_1. In general, λ_1 is directly proportional to the phase noise bandwidth (see Section 3.3.3). As this bandwidth increases, for example in the PARMA model by varying R_s, λ_1 increases and, hence, the second terms in (3.34) and (3.36) increase, thereby, resulting in a larger reduction from the fixed channel and Rayleigh fading capacities.

Equations (3.34) and (3.36) represents the capacity for a particular (j)th subcarrier. The net throughput of the system is obtained by summing the capacities over all subcarriers and dividing the result over the OFDM symbol duration which is $(N_\mathrm{c}+N_\mathrm{cp})T_\mathrm{s}$, where N_cp is the cyclic prefix length

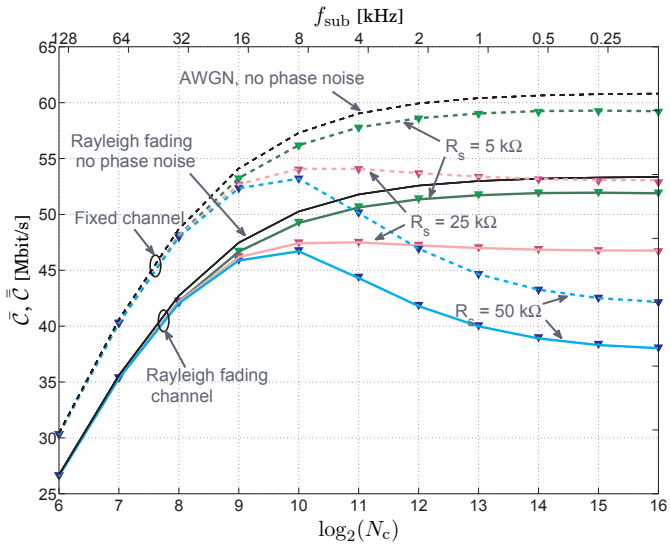


Figure 3.4. Net throughput plots of (3.37) for different phase noise PSD bandwidths obtained by varying the loop filter resistance R_s. The markers and solid/dashed lines represent the simulations and analytical plots, respectively. The PARMA filter parameters are given in Publication II.

and T_s is the sampling period. For the fixed channel case, we consider the realization wherein $\{\gamma_j\}_{j=0}^{N_c-1}$ are equal and, for the fading case, $\{\bar{\gamma}_j\}_{j=0}^{N_c-1}$ are set to the same average SNR. The net throughput is given by

$$\bar{\mathcal{C}} = \sum_{j=0}^{N_c-1} \frac{\bar{\mathcal{C}}_j}{(N_c+N_{cp})T_s} = \eta F_s \bar{\mathcal{C}}_j,\ \bar{\bar{\mathcal{C}}} = \sum_{j=0}^{N_c-1} \frac{\bar{\bar{\mathcal{C}}}_j}{(N_c+N_{cp})T_s} = \eta F_s \bar{\bar{\mathcal{C}}}_j, \quad (3.37)$$

where $F_s = 1/T_s$ is sampling frequency and $\eta = N_c/(N_c+N_{cp})$ is a measure of the loss in efficiency due to cyclic prefix.

Figure 3.4 shows the net throughput as a function of N_c for a fixed $N_{cp} = 64$. In the absence of phase noise, $\bar{\bar{\mathcal{C}}}$ and $\bar{\mathcal{C}}$ (shown by the black curves) increase with N_c and saturates to a particular value for $N_c \gg 1$. This is because $\eta \to 1$ for $N_c \gg N_{cp}$, i.e., the efficiency can be improved by choosing a large value of N_c in comparison with N_{cp}. However, in the presence of phase noise, we see that there is an optimal N_c for which $\bar{\mathcal{C}}$ and $\bar{\bar{\mathcal{C}}}$ are maximum. This can be explained as follows: For a fixed phase noise bandwidth, we know that $\bar{\mathcal{C}}_j$ and $\bar{\bar{\mathcal{C}}}_j$ decrease when N_c is increased. Thus, we have two conflicting scenarios, where $\bar{\mathcal{C}}$ and $\bar{\bar{\mathcal{C}}}$ increase with η and decrease with $\bar{\mathcal{C}}_j$ and $\bar{\bar{\mathcal{C}}}_j$, simultaneously. Thus, we could expect $\bar{\mathcal{C}}$ and $\bar{\bar{\mathcal{C}}}$ to increase with N_c up to a maximum as long as f_{sub} is large enough to cause the ICI to be small. Beyond this maximum, if f_{sub} is decreased then the resulting ICI causes $\bar{\mathcal{C}}_j$ and $\bar{\bar{\mathcal{C}}}_j$ to decrease much faster compared with the increase in η as evidenced in Fig. 3.4.

3.4 Characterization and Analysis Including Frequency Offset

The work in Publication III extends the analysis of Publications I and II to include the effect of carrier frequency offset. In this section, we summarize the results of Publication III.

The goal is to obtain analytical expressions of average capacity when the OFDM radio link is perturbed by both phase noise and carrier frequency offset. The starting point is the instantaneous SINR which is given by

$$\Upsilon_j = \frac{1-Y}{Y + \frac{\sigma_w^2}{\sigma_s^2 G_j}}, \tag{3.38}$$

where $Y = \sum_{k=1}^{N_c-1} |\delta_k|^2$ is the ICI power and $G_j = |\mathcal{H}_j|^2$. The coefficients δ_k denote the DFT of the combined transmitter and receiver phase noise along with frequency offset and is given by

$$\delta_k = \frac{1}{N_c} \sum_{n=0}^{N_c-1} e^{\jmath(\theta[n] + 2\pi \frac{f_\Delta}{N_c} n)} e^{-\jmath(2\pi kn)/N_c}, \tag{3.39}$$

where $\theta[n] = \theta^T[n] + \theta^R[n]$ and $f_\Delta = \frac{f_{\text{off}}}{f_{\text{sub}}}$; f_{off} denotes the frequency offset. From the definition of f_Δ, it is the normalized carrier frequency offset, where the normalization is w.r.t. f_{sub}. The SINR expression of (3.38) is of the exact same form as (3.7) which is the SINR with only phase noise. These are, however, different because now Y also incorporates information about the frequency offset.

3.4.1 Characterization of ICI Power

A suitable characterization of the ICI power is desired such that it facilitates evaluation of its PDF. Assuming a slow-varying phase noise process, such a characterization can be obtained using a Taylor series approximation and is given by

$$Y \approx d + \mathbf{x}^{\mathrm{T}} \left(\mathbf{a} + \mathbf{B}\mathbf{x} \right), \tag{3.40}$$

where $\mathbf{x}$ is an $N_{\mathrm{c}}(N_{\mathrm{c}}-1)/2$ dimensional Gaussian random vector with elements $\Delta\theta[i,l] = \theta[i+l-1] - \theta[l-1]$ whose variance $\sigma^2[i,l] = i\sigma^2$ for $i = 1, 2, \ldots, N_{\mathrm{c}} - l$ and $l = 1, 2, \ldots, N_{\mathrm{c}} - 1$. The respective diagonal elements of the diagonal matrix $\mathbf{B}$ and the elements of the column vector $\mathbf{a}$ are $\frac{\cos(\phi_i)}{N_c^2}$ and $\frac{2\sin(\phi_i)}{N_c^2}$ for $i = 1, 2, \ldots, N_{\mathrm{c}} - l$ and $l = 1, 2, \ldots, N_{\mathrm{c}} - 1$, where $\phi_i = \left(2\pi \frac{f_\Delta}{N_{\mathrm{c}}} i \right)$. The constant d is given by

$$d = 1 - \frac{1}{N_{\mathrm{c}}} \left[1 + \frac{2}{N_{\mathrm{c}}} \sum_{l=1}^{N_{\mathrm{c}}-1} \sum_{i=1}^{N_{\mathrm{c}}-l} \cos(\phi_i) \right]. \tag{3.41}$$

From (3.40), we see that the ICI power is given by the sum of Gaussian and gamma distributed random variables. In the absence of frequency offset, we have $\mathbf{a} = \mathbf{0}$ and $\mathbf{B}$ is the identity matrix, and thus, Y reduces to (3.11) whose PDF is given by (3.16).

3.4.2 PDF of Sum of Gamma and Gaussian Variates; Type I

The ICI power of (3.40) is given as a sum of correlated gamma and Gaussian distributed random variables. The diagonal elements of the diagonal matrix $\mathbf{B}$ take the form of $\cos(\phi_i)$ which depending upon the value of the f_Δ can result in a positive, negative or zero value. Enforcing the restriction $f_\Delta \leq \frac{1}{4}$ ensures that $\mathbf{B}$ is of full rank with positive diagonal values. With this assumption for the matrix $\mathbf{B}$, the PDF of Y can be derived on similar lines, as done for Theorem 3.3.1, by using the Moschopoulos technique, however, with some modifications [95]. We now have the following theorem[1]:

Theorem 3.4.1. *Let $Y = d + \mathbf{x}^{\mathrm{T}}(\mathbf{a} + \mathbf{B}\mathbf{x})$. Assume the diagonal matrix $\mathbf{B}$ to be of full rank with positive diagonal elements. Denote by $\mathbf{M}_\mathbf{x}$, of rank R, as the covariance matrix of the N-dimensional Gaussian random vector $\mathbf{x}$. The PDF of Y is given by*

$$p_Y(y) = K \sum_{k=0}^{\infty} \frac{\zeta_k (y-\mu)^{\frac{R}{2}+k-1} e^{\frac{-(y-\mu)}{\lambda_1}}}{\lambda_1^{\frac{R}{2}+k}\Gamma(\frac{R}{2}+k)} U(y-\mu), \tag{3.42}$$

where $U(y)$ is the unit step function. The coefficients ζ_k are obtained recursively as follows:

$$\zeta_0 = 1,\ \zeta_{k+1} = \frac{0.5}{k+1} \sum_{i=1}^{k+1} \left[\sum_{j=1}^{R} (1-\lambda_1/\lambda_j)^i \left(1 + \frac{i b_j (\lambda_1/\lambda_j)}{(1-\lambda_1/\lambda_j)}\right)\right] \zeta_{k+1-i}, \tag{3.43}$$

where $\{\lambda_i\}_{i=1}^R$ are the ordered non-zero eigenvalues (λ_1 being the minimum) of the matrix $2\Delta^{1/2}\mathbf{P}\mathbf{B}\mathbf{P}^{\mathrm{T}}\Delta^{1/2}$ with eigenvalue decomposition $\mathbf{V}\mathbf{\Lambda}\mathbf{V}^{\mathrm{T}}$. The vector $\mathbf{c} = \left(\mathbf{V}^{\mathrm{T}}\Delta^{1/2}\mathbf{P}\mathbf{a}\right)$ whose elements are denoted by c_i, and $b_i = c_i^2/\lambda_i^2$. The constant $K = K_{\mathrm{p}}K_{\mathrm{c}}$: $K_{\mathrm{p}} = \prod_{i=1}^{R}\left(\frac{\lambda_1}{\lambda_i}\right)^{\frac{1}{2}}$; and $K_{\mathrm{c}} = e^{-\frac{1}{2}\sum_{i=1}^{R} b_i}$. The delay factor $\mu = d - \tau$, where $\tau = \frac{1}{2}\sum_{i=1}^{R} b_i\lambda_i$ and $\mathbf{P} = \left[\mathbf{I}_R \ \ \mathbf{0}_{R\times(N-R)}\right]\mathbf{C}^{\mathrm{T}}$. The matrix $\mathbf{C}$ is obtained from $\mathbf{M}_\mathbf{x} = \mathbf{C}\mathbf{\Sigma}\mathbf{C}^{\mathrm{T}}$ which is an eigendecomposition of $\mathbf{M}_\mathbf{x}$. The matrix $\mathbf{\Delta}$ is diagonal whose elements are the non-zero eigenvalues of $\mathbf{M}_\mathbf{x}$.

[1] In Publication IV, a general result on the PDF of $Y = \mathbf{x}^{\mathrm{T}}(\mathbf{a} + \mathbf{B}\mathbf{x})$ is derived, where the restrictions of full-rank $\mathbf{B}$ and positive diagonal elements are removed. This general case is referred to as Type II and is summarized in Theorem 3.5.1.

Proof. A short version of the proof is given in Publication III. An elaborate version is given in Appendix A. □

In Fig. 3.5, we plot some PDF curves of the ICI power under the influence of carrier frequency offset and phase noise. The phase noise model used to generate these curves is the Wiener model. We consider slow and fast varying phase noise cases shown, respectively, by Figs. 3.5a and 3.5b. First, we observe that there is good agreement between analytical PDF of the ICI power of (3.42) and the simulated PDF shown by the marker lines. From Fig. 3.5a, for the slow varying case, we observe that the ICI power is more sensitive to the frequency offset demonstrated by the PDF getting broader and shifting to higher values of magnitude while, for the fast-varying case, the PDF is predominantly dictated by phase noise and seems insensitive to frequency offset. This behavior of the ICI power can also be inferred analytically using (3.42) whose behavior is mainly dictated by the parameter λ_1. We shall defer this analysis to the next section on average capacity, where we shall see a similar pattern.

3.4.3 Average Capacity

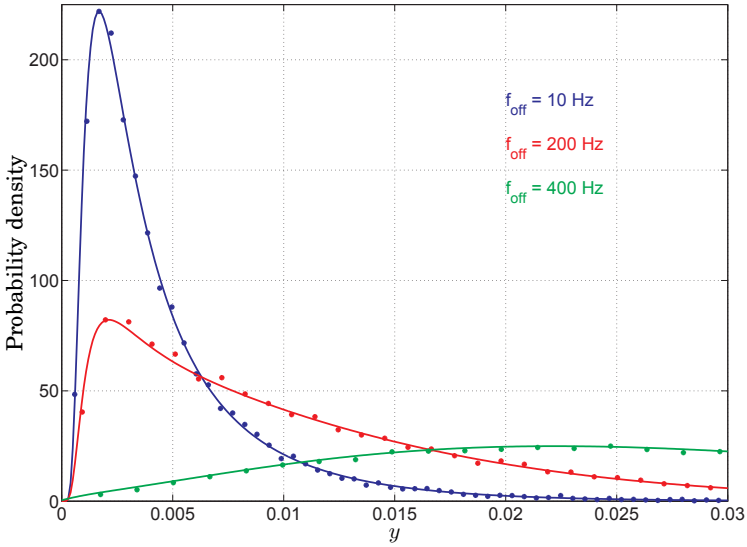
The average capacity is derived on similar lines as done for the phase noise-only case. The instantaneous capacity is given by $\mathcal{C}_j = \log_2(1+\Upsilon_j)$ which depends on the independent random variables Y and G_j. The average capacity is obtained by sequentially averaging $\mathcal{C}_j$ over the distributions of Y and G_j. Assuming a Rayleigh fading channel, the average capacity is derived in Publication III, and the final expression is given by

$$\bar{\bar{\mathcal{C}}}_j = \log_2(e)\left[e^{\frac{1}{\bar{\gamma}_j}}E_1\Big(\frac{1}{\bar{\gamma}_j}\Big) - K\sum_{k=0}^{\infty}\zeta_k e^{\frac{1}{r_k\bar{\gamma}_j}}E_1\Big(\frac{1}{r_k\bar{\gamma}_j}\Big)\right], \tag{3.44}$$

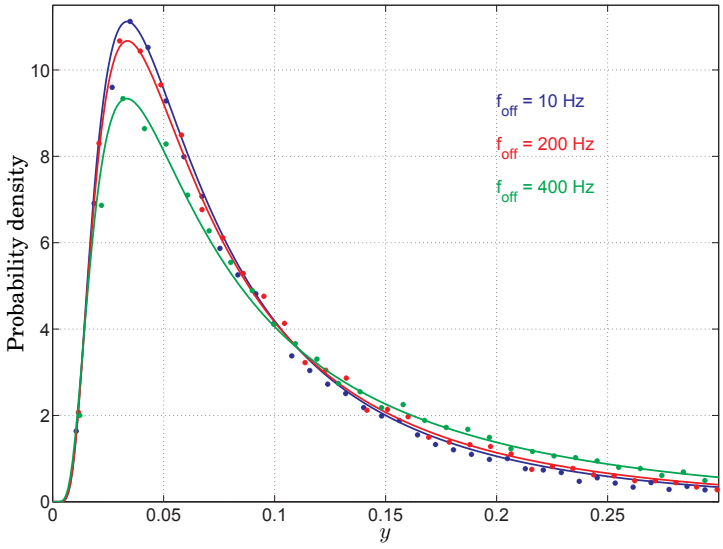
where $E_1(\cdot)$ is the exponential integral function of order one, $\bar{\gamma}_j = \frac{\bar{g}_j\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{w}}^2}$ denotes the average SNR and

$$r_k = \mu + \frac{\Gamma(R/2+k+1)\lambda_1}{\Gamma(R/2+k)}. \tag{3.45}$$

The first term in (3.44) represents the impairment-free channel capacity in a Rayleigh fading channel, while the second term results from the presence of phase noise and frequency offset, and the overall result is a reduction from the impairment-free case. The dependence of $\bar{\bar{\mathcal{C}}}_j$ on phase noise and frequency offset is mainly through the parameter λ_1. An increase in the phase noise bandwidth and frequency offset value causes an increases in λ_1 which results in a reduction of the capacity.



(a) Slow-varying phase noise case: $f_{3\text{dB}} = 10$ Hz.



(b) Fast-varying phase noise case: $f_{3\text{dB}} = 200$ Hz.

Figure 3.5. Comparison between analytical and simulated PDFs of Y. The markers and solid lines represent the simulations and analytical plots, respectively. System bandwidth is 9.14 MHz with $N_{\text{c}} = 2048$ and $f_{\text{sub}} = 4.5$ kHz.

Figure 3.6 shows average capacity plots as a function of the average SNR. The average capacity in the absence of phase noise and frequency offset is also shown for the purpose of comparison which is given by the black solid curve. This corresponds to the first term in (3.44). A general conclusion that can be made from the figure is that there is a net reduction of the average capacity in presence of phase noise and frequency

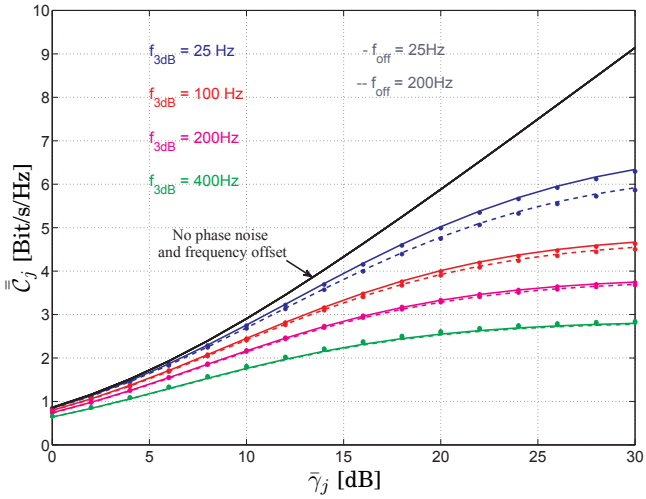
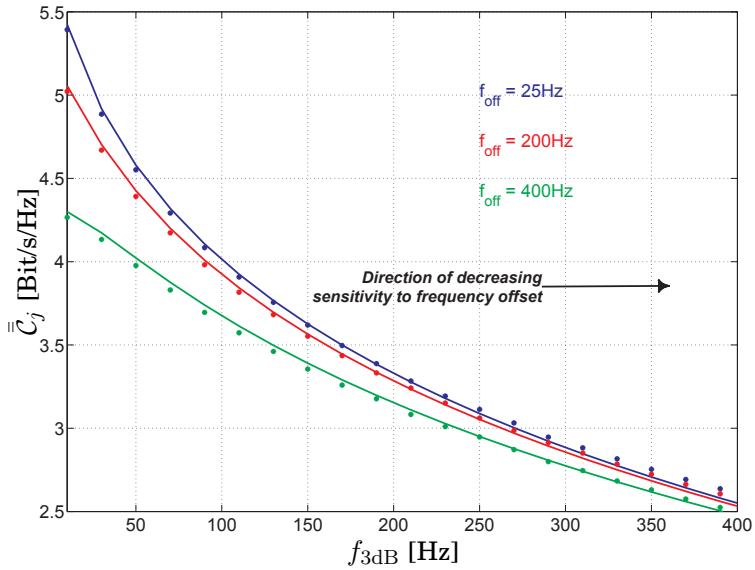


Figure 3.6. Comparison between simulated and analytical capacity plots of $\bar{\bar{C}}_j$ as a function of $\bar{\gamma}_j$ for different phase noise 3dB bandwidths $f_{3\mathrm{dB}}$ and carrier frequency offsets f_{off}. The markers and solid/dashed lines represent the simulations and analytical plots, respectively. System bandwidth is 9.14 MHz with $N_{\mathrm{c}} = 2048$ and $f_{\mathrm{sub}} = 4.5$ kHz. Channel is Rayleigh fading with 50 taps following an exponential power-delay profile of coherence bandwidth 400 kHz.

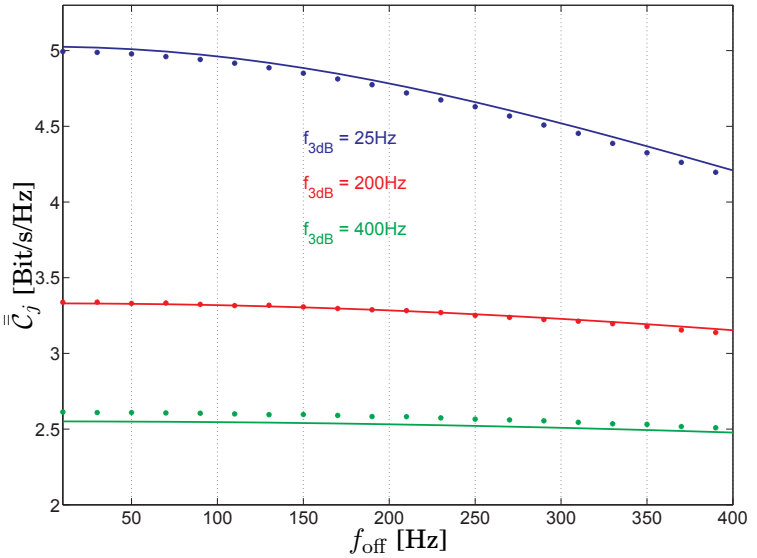
offset. This behavior can be analyzed analytically using (3.44) which is dependent upon the parameter λ_1. In Publication III, for a Wiener phase noise process, it is shown that $\lambda_1 \propto f_{3\mathrm{dB}}$. Thus, for a fixed value of the frequency offset, as $f_{3\mathrm{dB}}$ increases, so does λ_1 which causes the second term in (3.44) to increase.

Figure 3.7 shows the sensitivity of the capacity to phase noise and frequency offset. Specifically, in Fig. 3.7a, $\bar{\bar{C}}_j$ is plotted as a function of $f_{3\mathrm{dB}}$ for different values of f_{off}. As seen from the figure, for small values of $f_{3\mathrm{dB}}$, the capacity is more sensitive to frequency offset when compared to the case with high values of $f_{3\mathrm{dB}}$. This behavior can also be seen in Fig. 3.7b where, for $f_{3\mathrm{dB}} = 400$ Hz, $\bar{\bar{C}}_j$ is practically insensitive to the carrier frequency offset.

The insensitivity of the capacity to frequency offset at high values of $f_{3\mathrm{dB}}$ can also be inferred analytically. The diagonal matrix $\mathbf{B}$ is composed of $N_{\mathrm{c}} - 1$ basic elements $\frac{\cos(\phi_i)}{N_c^2}$ for $i = 1, 2, \ldots, N_{\mathrm{c}} - 1$, where $\phi_i = 2\pi \frac{f_\Delta}{N_{\mathrm{c}}} i$. The parameter λ_1 is the smallest eigenvalue of the matrix $2\Delta^{1/2}\mathbf{P}\mathbf{B}\mathbf{P}^{\mathrm{T}}\Delta^{1/2}$. Thus, in addition to $\lambda_1 \propto f_{3\mathrm{dB}}$, we also have that $\lambda_1 \propto \cos\left(\frac{f_{\mathrm{off}}}{f_{\mathrm{sub}}}\right)$. However, fast-varying phase noise processes, the linear dependency of λ_1 on $f_{3\mathrm{dB}}$ has a larger effect than the non-linear dependency on f_{off} through the cosine function whose range is limited between minus one and plus one.



(a) Average capacity as a function of phase noise 3dB bandwidth for $\bar{\gamma}_j = 20$ dB.



(b) Average capacity as a function of frequency offset for $\bar{\gamma}_j = 20$ dB.

Figure 3.7. Sensitivity of average capacity to phase noise and frequency offset. The markers and solid lines represent the simulations and analytical plots, respectively.

Thus, λ_1 and, hence, $\bar{\bar{C}}_j$ in (3.44) , are practically insensitive with respect to changes in the frequency offset and their behavior is mainly dictated by phase noise.

3.5 Characterization of Phase Noise Spectral Components

In the previous sections, characterization of the ICI power in terms of a sum of gamma random variables facilitated in evaluation of the average capacity. In this section, we seek a similar characterization of the components of $\boldsymbol{\delta}$ in terms of well-defined elements. In this thesis, we refer to components of $\boldsymbol{\delta}$ as 'phase noise spectral components'. Such a characterization is useful from the perspective of phase noise estimation, where one way of removing phase noise is to first estimate $\boldsymbol{\delta}$ and then perform compensation.

In Publication IV, the PDF of the real and imaginary parts of δ_k are derived which is then used in deriving Bayesian minimum mean square error (MMSE) estimates. Although such an approach of estimating each component separately is inefficient, the characterization and the resulting PDF nevertheless point towards some interesting aspects.

3.5.1 Taylor Series Approximation of Spectral Components

The components $\boldsymbol{\delta}_k$ are given by

$$\delta_k = \frac{1}{N_c} \sum_{n=0}^{N_c-1} e^{\jmath\theta[n]} e^{-\jmath(2\pi kn)/N_c}, \tag{3.46}$$

where $\theta[n]$ represents the phase noise. We can also include the effect of carrier frequency offset, however, without any loss in generality, we only consider phase noise. Taking the real and imaginary parts of (3.46), and performing a second-order Taylor-series expansion while assuming that the phase noise process is slow-varying, we arrive at

$$\delta_k^{\mathrm{r}} \approx d_k^{\mathrm{r}} - \mathbf{x}^{\mathrm{T}} \left(\mathbf{a}_k^{\mathrm{r}} + \mathbf{B}_k^{\mathrm{r}}\right) \mathbf{x}, \;\; \delta_k^{\mathrm{i}} \approx d_k^{\mathrm{i}} + \mathbf{x}^{\mathrm{T}} \left(\mathbf{a}_k^{\mathrm{i}} - \mathbf{B}_k^{\mathrm{i}}\right) \mathbf{x}, \tag{3.47}$$

where the respective δ_k^{r} and δ_k^{i} are the real and imaginary parts of δ_k; and $\mathbf{x} = [\theta[0],\ \theta[1], \ldots, \theta[N_{\mathrm{c}}-1]]^{\mathrm{T}}$. The vectors $\mathbf{a}_k^{\mathrm{r}}$ and $\mathbf{a}_k^{\mathrm{i}}$ are given by

$$\mathbf{a}_k^{\mathrm{r}} = \frac{1}{N_{\mathrm{c}}} \left[0, \sin\left(\frac{-2\pi k}{N_{\mathrm{c}}}\right), \sin\left(\frac{-4\pi k}{N_{\mathrm{c}}}\right), \ldots, \sin\left(\frac{-2\pi k(N_{\mathrm{c}}-1)}{N_{\mathrm{c}}}\right)\right]^{\mathrm{T}}; \tag{3.48}$$

$$\mathbf{a}_k^{\mathrm{i}} = \frac{1}{N_{\mathrm{c}}} \left[1, \cos\left(\frac{-2\pi k}{N_{\mathrm{c}}}\right), \cos\left(\frac{-4\pi k}{N_{\mathrm{c}}}\right), \ldots, \cos\left(\frac{-2\pi k(N_{\mathrm{c}}-1)}{N_{\mathrm{c}}}\right)\right]^{\mathrm{T}}. \tag{3.49}$$

The matrices $\mathbf{B}_k^{\mathrm{r}} = \frac{1}{2}\operatorname{diag}\left(\mathbf{a}_k^{\mathrm{i}}\right)$ and $\mathbf{B}_k^{\mathrm{i}} = \frac{1}{2}\operatorname{diag}\left(\mathbf{a}_k^{\mathrm{r}}\right)$. The constants

$$d_k^{\mathrm{r}} = \sum_{l=0}^{N_{\mathrm{c}}-1} \mathbf{a}_{lk}^{\mathrm{i}};\ d_k^{\mathrm{i}} = \sum_{l=0}^{N_{\mathrm{c}}-1} \mathbf{a}_{lk}^{\mathrm{r}}. \tag{3.50}$$

From (3.47), we see that, for slow-varying phase noise processes, the real and imaginary parts of δ_k are given by the sum of Gaussian and gamma random variables, where $\mathbf{x}$ is assumed to be Gaussian distributed.

3.5.2 PDF of Sum of Gamma and Gaussian Variates; Type II

In Publication III, the PDF of $Y = d + \mathbf{x}^{\mathrm{T}}(\mathbf{a} + \mathbf{B}\mathbf{x})$ is derived under the assumption of full-rank diagonal $\mathbf{B}$ matrix with positive diagonal values. The result is also summarized in Theorem 3.4.1 of Section 3.4.2. However, the theorem is not applicable to (3.47) since the diagonal matrices $\mathbf{B}_k^{\mathrm{r}}$ and $\mathbf{B}_k^{\mathrm{i}}$ are rank deficient as well as harboring negative diagonal values. A similar approach to that used in Theorem 3.4.1 can be used to derive the PDF, but it must be modified to incorporate aspects of negativity and rank deficiency. The following theorem summarizes the general case:

Theorem 3.5.1. *Let $Y = \mathbf{x}^{\mathrm{T}}(\mathbf{a} + \mathbf{B}\mathbf{x})$, where $\mathbf{x}$ is an N-dimensional zero mean Gaussian random vector with covariance matrix $\mathbf{M}_{\mathrm{x}}$ of rank R_{x}, $\mathbf{a}$ is a column vector and $\mathbf{B}$ is a real diagonal matrix of any rank. The PDF of Y can be equivalently expressed as a convolution of a Gaussian distributed PDF and a weighted sum of gamma distributed PDFs as shown below*

$$
\begin{aligned}
&P_Y(y+\tau) = P_{\mathcal{G}}(y) \star P_{\mathcal{N}}(y), \\
&= K \left[\sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \eta_k \zeta_j \left(\sum_{l=1}^{R_k} A_l^{kj} \mathcal{G}(y; l, \beta_1) U(y) + \sum_{m=1}^{L_j} \tilde{A}_m^{kj} \mathcal{G}(-y; m, |\gamma_1|) U(-y) \right) \right] \\
&\qquad \star \left[\frac{1}{\sqrt{2\pi\sigma_{\mathcal{N}}^2}} e^{\frac{-y^2}{2\sigma_{\mathcal{N}}^2}} \right], \qquad (3.51)
\end{aligned}
$$

where $\star$ denotes linear convolution. The term $\mathcal{G}(y; k, \theta)$ denotes a gamma distribution with shape parameter k and scale parameter θ. The step function is denoted by $U(y)$. The parameters β_i and γ_i are the R positive and L negative eigenvalues of the matrix $2\Delta^{1/2}\mathbf{P}\mathbf{B}\mathbf{P}^{\mathrm{T}}\Delta^{1/2}$. β_1 is the minimum among β_i, while γ_1 is the maximum among γ_i. The $\mathbf{P}$ and $\mathbf{\Delta}$ matrices are obtained from the eigenvalue decomposition of $\mathbf{M}_{\mathbf{x}}$ as follows:

$$
\mathbf{M}_{\mathbf{x}} = \mathbf{C} \begin{pmatrix} \mathbf{\Delta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{C}^{\mathrm{T}}; \qquad (3.52)
$$

$$
\mathbf{P} = \begin{bmatrix} \mathbf{I}_{R_{\mathrm{x}}} & \mathbf{0}_{R_{\mathrm{x}} \times (N - R_{\mathrm{x}})} \end{bmatrix} \mathbf{C}^{\mathrm{T}}, \qquad (3.53)
$$

where the $R_{\mathrm{x}} \times R_{\mathrm{x}}$ $\mathbf{\Delta}$ matrix is diagonal and consists of the non-zero eigenvalues of $\mathbf{M}_{\mathbf{x}}$. The coefficients in (3.51) *are given as follows:*

$$
R_k = \frac{R}{2} + k, \; L_j = \frac{L}{2} + j, \; \eta_0 = 1, \; \zeta_0 = 1; \qquad (3.54)
$$

$$\eta_{k+1} = \frac{0.5}{k+1} \sum_{i=1}^{k+1} \left(\sum_{j=1}^{R} \mathcal{Y}(i, \beta_1, \beta_j, u_j) \right) \eta_{k+1-i}; \tag{3.55}$$

$$\zeta_{k+1} = \frac{0.5}{k+1} \sum_{i=1}^{k+1} \left(\sum_{j=1}^{L} \mathcal{Y}(i, \gamma_1, \gamma_j, v_j) \right) \zeta_{k+1-i}; \tag{3.56}$$

$$\mathcal{Y}(i, x, y, r) = \left(1 - \frac{x}{y}\right)^i \left[1 + \frac{ir}{\frac{y}{x} - 1}\right]; \tag{3.57}$$

$$A_l^{kj} = \frac{(R_k + L_j - l - 1)!(\frac{1}{\beta_1} + \frac{1}{|\gamma_1|})^{-(R_k + L_j - l)}}{\beta_1^{(R_k - l)} |\gamma_1|^{L_j} (R_k - l)!(L_j - 1)!}; \tag{3.58}$$

$$\tilde{A}_m^{kj} = \frac{(R_k + L_j - m - 1)!(\frac{1}{\beta_1} + \frac{1}{|\gamma_1|})^{-(R_k + L_j - m)}}{\beta_1^{(R_k)} |\gamma_1|^{(L_j - m)} (L_j - m)!(R_k - 1)!}. \tag{3.59}$$

The constant K is given by

$$K = K_c \prod_{i=1}^{R} \left(\frac{\beta_1}{\beta_i}\right)^{\frac{1}{2}} \prod_{i=1}^{L} \left(\frac{\gamma_1}{\gamma_i}\right)^{\frac{1}{2}}, \tag{3.60}$$

where $K_c = e^{-\frac{1}{2}\left(\sum_{i=1}^{R} u_j + \sum_{i=1}^{L} v_j\right)}$. The elements u_j and v_j are obtained as follows: Denote the eigendecomposition of $2\Delta^{1/2}\mathbf{P}\mathbf{B}\mathbf{P}^{\mathrm{T}}\Delta^{1/2}$ by $\mathbf{V}\mathbf{\Lambda}\mathbf{V}^{\mathrm{T}}$, where λ_i are the diagonal values of $\mathbf{\Lambda}$. Denote the respective $\mathcal{R}$, $\mathcal{L}$ and $\mathcal{Z}$ as the index of positive, negative and zero eigenvalues of $2\Delta^{1/2}\mathbf{P}\mathbf{B}\mathbf{P}^{\mathrm{T}}\Delta^{1/2}$. Define the vector $\mathbf{c} = \mathbf{V}^{\mathrm{T}}\Delta^{1/2}\mathbf{P}\mathbf{a}$ whose elements are denoted by c_k. We have

$$\tau = \frac{1}{2}\left(\sum_{k \in \mathcal{R}} \left(\frac{c_k^2}{\lambda_k}\right) + \sum_{k \in \mathcal{L}} \left(\frac{c_k^2}{\lambda_k}\right) \right), \tag{3.61}$$

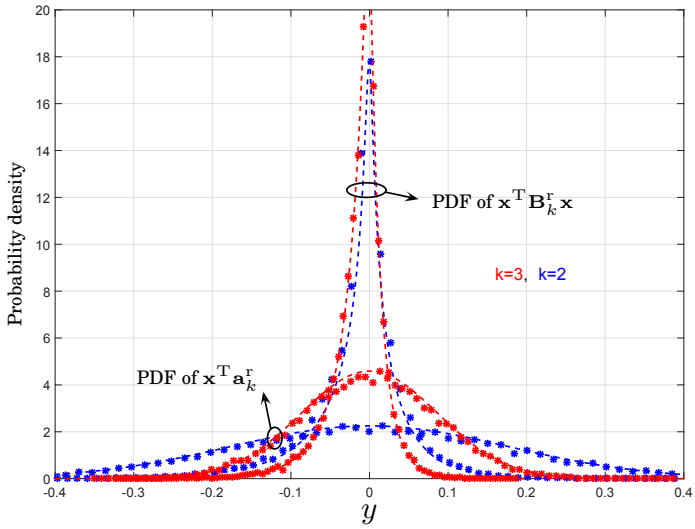
$$u_j = (c_k^2/\lambda_k^2) \text{ for } k \in \mathcal{R}, j = 1, 2, \ldots, R, \tag{3.62}$$

$$v_j = (c_l^2/\lambda_l^2) \text{ for } l \in \mathcal{L}, j = 1, 2, \ldots, L. \tag{3.63}$$

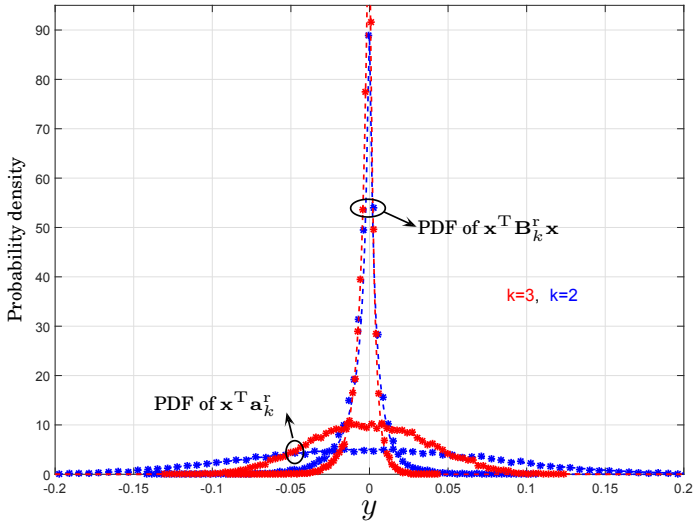
Finally, the variance of the Gaussian distribution $P_{\mathcal{N}}(y)$ is given by $\sigma_{\mathcal{N}}^2 = \sum_{k \in \mathcal{Z}} c_k^2$.

Proof. For a short-version of the proof, see Publication IV. For the complete version, see Appendix A.2. □

Figures 3.8 and 3.9 show the PDF of the linear and quadratic terms that comprise the real and imaginary parts of δ_k. The analytical PDFs are obtained by making use of (3.51). For example, the PDF of $\mathbf{x}^{\mathrm{T}}\mathbf{a}_k^{\mathrm{r}}$ is obtained by setting $\mathbf{B} = \mathbf{0}$ in Theorem 3.5.1. Similarly, PDF of $\mathbf{x}^{\mathrm{T}}\mathbf{B}_k^{\mathrm{r}}\mathbf{x}$ is obtained by setting the vector $\mathbf{a} = \mathbf{0}$. From Figs. 3.8 and 3.9, for the set phase noise bandwidth, we see good agreement between the PDF predicted by (3.51) and the simulated PDF. Also observed from the figure is that the linear Gaussian terms comprising both δ_k^{r} and δ_k^{i} is the stronger component compared to the quadratic and gamma-like distributed terms.

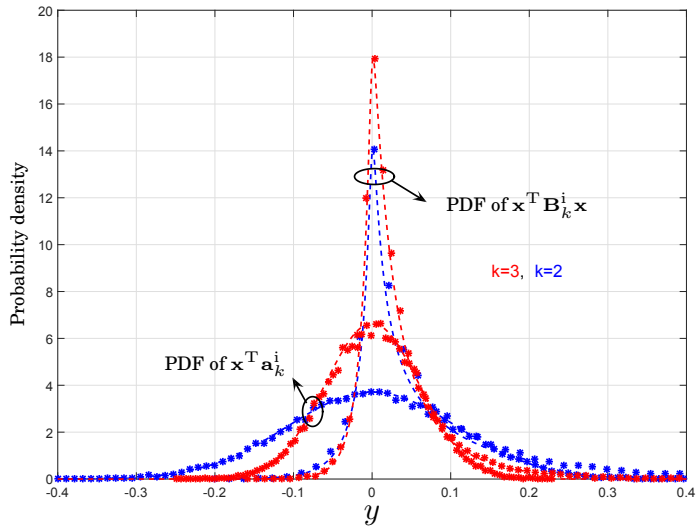


(a) PDF of the linear and quadratic terms constituting the real part of δ_k for $f_{3\mathrm{dB}}/f_{\mathrm{sub}} = 0.06$.

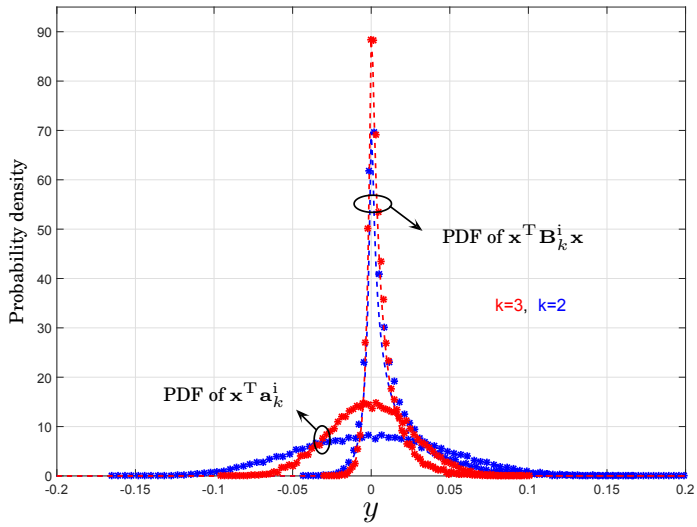


(b) PDF of the linear and quadratic terms constituting the real part of δ_k for $f_{3\mathrm{dB}}/f_{\mathrm{sub}} = 0.01$.

Figure 3.8. Comparison between analytical and simulated PDFs of the Gaussian and gamma variables of (3.47) for $k = 2$ and $k = 3$ with $N_{\mathrm{c}} = 16$. Dashed lines are the analytical curves, while markers represent the simulated PDF. The Wiener phase noise 3-dB bandwidth and OFDM subcarrier spacing are denoted by $f_{3\mathrm{dB}}$ and f_{sub}, respectively.



(a) PDF of the linear and quadratic terms constituting the imaginary part of δ_k for $f_{3\mathrm{dB}}/f_{\mathrm{sub}} = 0.06$.



(b) PDF of the linear and quadratic terms constituting the imaginary part of δ_k for $f_{3\mathrm{dB}}/f_{\mathrm{sub}} = 0.01$.

Figure 3.9. Comparison between analytical and simulated PDFs of the Gaussian and gamma variables of (3.47) for $k = 2$ and $k = 3$ with $N_{\mathrm{c}} = 16$. Dashed lines are the analytical curves, while markers represent the simulated PDF. The Wiener phase noise 3-dB bandwidth and OFDM subcarrier spacing are denoted by $f_{3\mathrm{dB}}$ and f_{sub}, respectively.

3.6 Discussion

This chapter provides novel results on characterization and performance analysis of OFDM in the presence of both transmitter and receiver phase noise. Specifically, the aim of this chapter is to ascertain, both quantitatively and qualitatively, the degradation in capacity assuming a Gaussian input alphabet. Restricting the input to be drawn from a Gaussian distribution simplifies the analysis and highlights the degradation from the phase noise free case. In that sense, the capacity expressions presented in this chapter are not representative of the channel capacity in a strict information-theoretic sense.

The gaining popularity of OFDM in past two decades is positively correlated with studies, mainly by the research community, demonstrating the sensitivity of OFDM to phase noise and RF-impairments in general. All of these studies demonstrated degradation in the SINR, BEP and SEP in the presence of phase noise, where the overall progression of these studies was towards obtaining more accurate analytical expressions of the aforementioned performance metrics, thus, enabling the system designer as the oracle of reliable performance prediction. Channel capacity is another useful performance metric that indicates the data rate of the communication system. Prior to the work in Publications I, II and III, no studies on capacity of OFDM systems impaired by phase noise were available in the open literature. One of the objectives of this thesis is to fill this void.

The capacity expressions derived in this chapter can be beneficial to the RF system design engineer, where a cause-effect type of relationship between the designed oscillator or PLL-based device parameters and channel capacity can be seen.The capacity derived depends on the phase noise processes through the so-called ICI power which captures the total interference power caused by phase noise. The PDF of this ICI power for any Gaussian phase noise process is derived in this thesis and shown in this chapter to be a sum of gamma distributed random variables. The parameters of this PDF take on different values depending upon the type of phase noise process and the set phase noise level. Models for phase noise in free-running oscillators and PLL-based devices are well-established, wherein the model parameters relate in some non-linear fashion with the circuit design parameters. With these models in place, the ICI power PDF parameters can be numerically computed using which the capacity can be ascertained.

4. Estimation in OFDM Systems under Phase Noise

Phase noise estimation has become a necessary functionality that must be performed in order to obtain reliable estimates of the transmitted symbols. In order to perform phase noise estimation, reliable channel estimates are also desired. In certain cases, depending upon the OFDM system parameters and phase noise level, the channel estimation step may completely ignore the contribution of phase noise while, in other cases, the estimation process must incorporate the effect of phase noise.

In the first part of this chapter, we review some of the state-of-the-art methods for estimation in OFDM systems impaired by phase noise. Two approaches are typically used: *Isolated estimation* and *Joint estimation*. In the isolated approach, channel estimation, phase noise estimation and symbol estimation are separate functional blocks where each one performs the desired functionality while assuming that the dependent parameters can be obtained from the others. Such an approach may not be *statistically optimal*. In the joint approach, phase noise estimation or some knowledge of it is combined with channel estimation and symbol estimation such that they yield statistically optimal joint estimates.

The second part of this chapter summarizes the contributions of this thesis related to phase noise estimation. Specifically, two new phase noise estimation schemes proposed in Publications V and VI are reviewed. Before proceeding, we would like to remind the reader that there is extensive work on estimation in single-carrier systems impaired by phase noise. With the emergence of *massive MIMO (multiple-input multiple-output)*, there is renewed interest for phase noise estimation with the principal aim of seeking algorithms with reduced computational complexity. Some recent works can be found in [99–103] and references therein.

4.1 State-of-the-Art Estimation Schemes for OFDM

Most methods for phase noise estimation in OFDM fall in either categories of isolated or joint estimation. Both these approaches in a sense have a historical footprint wherein, initially, the methods developed were isolated estimation approaches while, only very recently, joint estimation approaches have become very popular.

Consider the OFDM system model impaired by only receiver phase noise which is given by (see Chapter 2.3.3 for details)

$$\mathbf{r} = \mathbf{VHs} + \mathbf{w}, \tag{4.1}$$

where the matrix $\mathbf{V}$ is column-wise circulant with the first column vector $\boldsymbol{\delta}$ given by

$$\delta_k = \sum_{n=0}^{N_c-1} \frac{e^{\jmath\theta[n]}}{N_c} e^{-\jmath(2\pi kn)/N_c}, k = 0, 1, \ldots, N_c - 1, \tag{4.2}$$

where $\theta[n]$ denotes the receiver phase noise. Here, the unknowns are the channel matrix described by $\mathbf{H}$, the phase noise matrix $\mathbf{V}$ and the desired transmitted symbol vector s. The unknowns can equivalently be described in terms of their time-domain counterparts.

Typically, to estimate the channel, a *preamble symbol* is used, where the symbol vector s is known to the receiver. In this thesis, we refer to this phase of transmission as the *pilot phase*. The *data phase* consists of regular transmission of symbols, and the task of the receiver is to recover s using the estimate of the channel obtained from the pilot phase. Such a transmission method is based on the assumption of a *quasi-static fading* channel, where the channel is assumed to be static for a certain length of the data phase. During both pilot and data phases, we have phase noise present in the radio link and, depending upon its severity, it can result in poor channel and symbol estimates.

The presence of phase noise necessitates estimating it or, at the very least, incorporating the information during channel and symbol estimation. By estimation of phase noise, we mean either estimating the spectral vector $\boldsymbol{\delta}$ or its time-domain equivalent or the actual phase noise realization itself, i.e., $\theta[n]$. Most approaches used in the general literature, related to phase noise estimation, fall in either category of isolated or joint approach. The basic ideology of these approaches is shown in Fig. 4.1. As seen in the figure, for all the approaches, channel estimation is performed in the preamble phase. These channel estimates are used in the data-phase for phase noise and symbol estimation. In the isolated approach,

phase noise is estimated independently of the channel estimation step and symbol estimation step while, in the *fully-joint* approach, phase noise is incorporated during channel and symbol estimation. In between these extremes is a *semi-joint* approach where phase noise is incorporated only for the channel estimation step, while the symbol estimation step ignores any kind of phase noise knowledge.

In isolated estimation approaches, the channel is estimated using methods that assume there is no phase noise present in the system. Depending upon the level of phase noise, poor channel estimates can be obtained which effectively will result in poor symbol estimates. The fully-joint and semi-joint approaches seek to alleviate this problem by taking phase noise into consideration during channel estimation. Such a joint approach will yield reliable channel estimates. Similarly, performing independently the phase noise estimation and symbol estimation steps, as done in the isolated and semi-joint approaches, may not necessarily be optimal in a statistical sense. Optimally statistical estimates can be obtained by jointly estimating the unknown parameters as done in the fully-joint approach. The way this is done is by using *Bayesian inference techniques* which are a set of methods that naturally allow for estimation of multiple parameters while ensuring some form of statistical optimality [104].

With this general classification in mind, we are now ready to review some of the state-of-the-art methods on isolated and joint approaches to phase noise estimation. The subject of phase noise estimation comes under the broader class of *synchronization* which includes also carrier frequency offset and timing offset [105]. Traditional signal processing methods for synchronization were mainly developed for single-carrier systems. These works have been extended to OFDM systems with some studies dating back to over twenty years ago [106]. However, the major explosion in phase noise estimation algorithms for OFDM has mainly coincided with the start of the new millennium. We shall, thus, mainly focus our attention on works developed since this period.

4.1.1 Separate Phase Noise and Symbol Estimation

In this section, we focus on phase noise estimation using the isolated approach wherein phase noise is first estimated and removed before performing symbol estimation/detection. Typically in these approaches, full channel knowledge is either assumed or some estimate of it is already available.

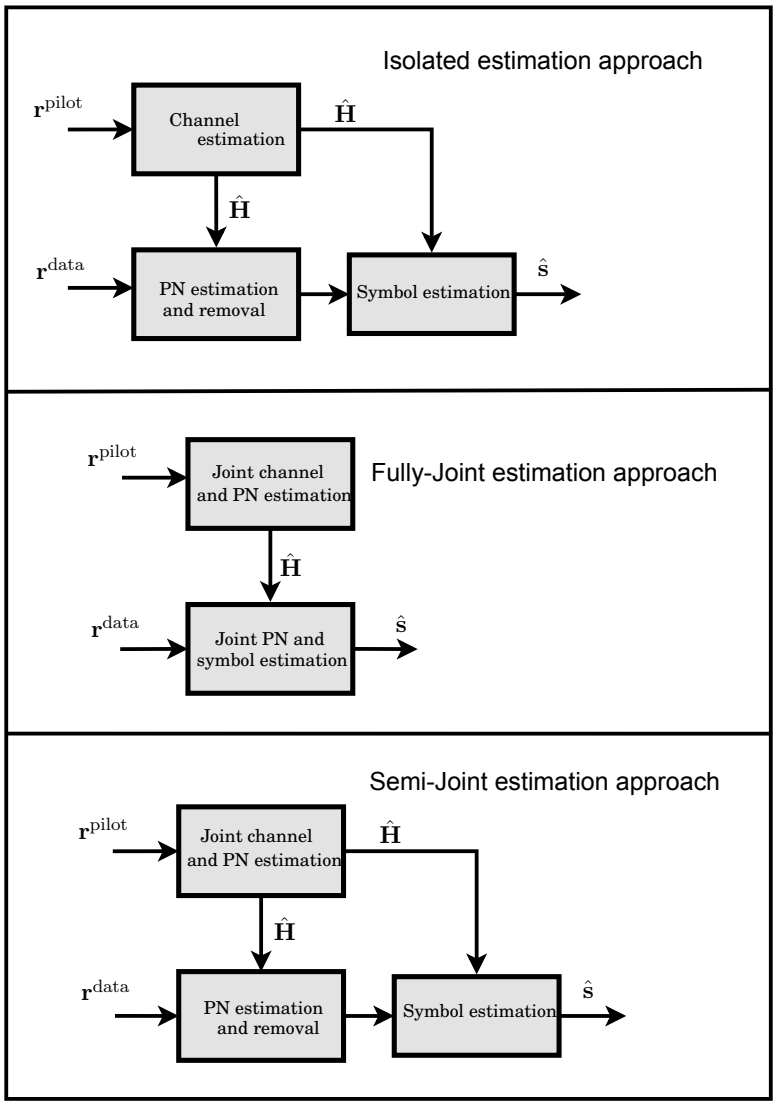


Figure 4.1. Isolated and joint phase noise estimation approaches. The respective $\mathbf{r}^{\text{pilot}}$ and $\mathbf{r}^{\text{data}}$ vectors denote the received OFDM symbol vector of (4.1) during the preamble phase and data phase.

Basis expansion approach of [107]

A very simple yet efficient method, based on exploiting the low-pass nature of phase noise processes, is explored in [107]. To understand the principle, consider (4.1). The maximum likelihood (ML) estimate of s in (4.1), assuming additive white Gaussian noise, is given by

$$\begin{aligned}\hat{\mathbf{s}} &= \hat{\mathbf{H}}^{-1}\mathbf{V}^{\dagger}\mathbf{r} \\ &= \hat{\mathbf{H}}^{-1}\mathbf{F}\,\mathrm{diag}\left(\mathbf{F}^{\dagger}\mathbf{r}\right)\boldsymbol{\phi}^{*} \\ &= \mathbf{D}\boldsymbol{\phi}^{*}, \end{aligned} \tag{4.3}$$

where $\mathbf{F}$ is the $N_c \times N_c$ DFT matrix,

$$\mathbf{D} = \hat{\mathbf{H}}^{-1}\mathbf{F}\,\mathrm{diag}\left(\mathbf{F}^{\dagger}\mathbf{r}\right) \tag{4.4}$$

and $\mathrm{diag}\,(\mathbf{x})$ denotes a diagonal matrix with $\mathbf{x}$ forming the main diagonal. The vector $\boldsymbol{\phi} = \left[e^{j\theta[0]},\ e^{j\theta[1]}, \ldots, e^{j\theta[N_c-1]}\right]^{\mathrm{T}}$ with $\boldsymbol{\phi}^*$ denoting the vector whose elements are the conjugate of the elements of $\boldsymbol{\phi}$. In arriving at (4.3), we used the fact that $\mathbf{V} = \mathbf{F}\,\mathrm{diag}\,(\boldsymbol{\phi})\,\mathbf{F}^{\dagger}$. The matrix $\hat{\mathbf{H}}$ denotes our estimate of $\mathbf{H}$. We see that an estimate of ϕ^* can be obtained if we knew the left-hand-side of (4.3).

In general, we do have access to certain components of $\mathbf{s}$ in the form of pilot symbols. Let $\mathbf{s}_{\mathcal{K}}$ denote the $P \times 1$ vector of pilot symbols, where $\mathcal{K}$ is the set of pilot indices. Using this symbol vector, an estimate of ϕ^* can be obtained as follows: In general, phase noise processes vary slowly over time effectively rendering them to be low pass in nature. Thus, the vector $\boldsymbol{\phi}^*$ can be modeled using few basis vectors as $\boldsymbol{\phi}^* = \mathbf{B}\beta$, where the $N_c \times N$ matrix $\mathbf{B}$ represents the basis set, and β represents the vector of associated weights. In [107], the authors propose to use a discrete Fourier basis set and a discrete cosine basis set. Assuming sufficient number of pilot symbols, i.e., $P > N$, the least-squares estimate of β is obtained as [108]

$$\begin{aligned}\hat{\beta} &= \arg\min_{\beta} \|\mathbf{s}_{\mathcal{K}} - \mathbf{D}_{\mathcal{K}}\mathbf{B}\beta\|^2 \\ &= \left(\mathbf{B}^{\dagger}\mathbf{D}_{\mathcal{K}}^{\dagger}\mathbf{D}_{\mathcal{K}}\mathbf{B}\right)^{-1}\mathbf{B}^{\dagger}\mathbf{D}_{\mathcal{K}}^{\dagger}\mathbf{s}_{\mathcal{K}},\end{aligned} \tag{4.5}$$

where $\mathbf{D}_{\mathcal{K}}$ is the matrix obtained after picking the rows of $\mathbf{D}$ that are indexed by the set $\mathcal{K}$.

Using (4.5), the estimate of $\boldsymbol{\phi}^*$ is given by $\hat{\boldsymbol{\phi}}^* = \mathbf{B}\hat{\beta}$. We finally obtain our estimate of the transmitted symbols by using $\hat{\boldsymbol{\phi}}^*$ in (4.3). By estimating $\boldsymbol{\phi}^*$ separately, and assuming an already available channel estimate $\hat{\mathbf{H}}$, we see that this phase noise estimation method falls under the isolated approach.

Basis expansion approach of [109] and [110]

The low pass nature of the phase noise process is also exploited in the works of [109] and [110] and in a manner similar to that of [107], i.e., ϕ is expressed using few basis vectors. Specifically, the DFT basis set is used in [109] and [110]. Expressing (4.1) in terms of the spectral vector $\boldsymbol{\delta} = \mathbf{F}\phi$, we have

$$\mathbf{r} = \mathbf{A}\boldsymbol{\delta} + \mathbf{w}, \tag{4.6}$$

where the first row of $\mathbf{A}$ is given by $[H_0 s_0,\ H_1 s_1, \ldots, H_{N_c-1} s_{N_c-1}]$, and the kth row is obtained by circularly left-shifting the first row $k-1$ times. Since ϕ is typically a low-pass process, it suffices to estimate only a few components of $\boldsymbol{\delta}$ as follows:

$$\mathbf{r} = \tilde{\mathbf{A}}\underline{\boldsymbol{\delta}} + \mathbf{v} + \mathbf{w}, \tag{4.7}$$

where $\underline{\boldsymbol{\delta}} = \left[\delta_0,\ \delta_1, \ldots, \delta_{N/2-1},\ \delta_{N_c-N/2},\ \delta_{N_c-(N/2+1)}, \ldots, \delta_{N_c-1}\right]^{\mathrm{T}}$ comprises of the low frequency components of ϕ. The $N_c \times N$ matrix $\tilde{\mathbf{A}}$ is obtained from $\mathbf{A}$ by picking out the N columns corresponding to the elements of $\underline{\boldsymbol{\delta}}$. The vector $\mathbf{v}$ corresponds to the unestimated part of $\boldsymbol{\delta}$.

In [109], the authors derive the ML estimate (MLE) and linear MMSE estimate of $\boldsymbol{\delta}$ based on the linear model of (4.6) while, in [110], the authors derive the MMSE estimate of $\underline{\boldsymbol{\delta}}$ using the model of (4.7). The derived estimators require knowledge of the matrices $\mathbf{A}$ and $\tilde{\mathbf{A}}$ which are composed of the channel frequency responses H_j and the transmitted symbols s_j. It is assumed that channel estimates and tentative decisions on s_j are available which are then used to form the required matrices and, finally to obtain a phase noise estimate. Using this estimate, the received signal is cleaned by removing the phase noise and, after performing channel equalization and symbol detection, the new symbol estimates are used to update $\tilde{\mathbf{A}}$ and the phase noise estimate. This process is repeated for a certain number of times.

Data-aided-based CPE estimation of [77] and [111]

The impact of phase noise on each component of $\mathbf{s}$ can be seen by explicitly writing out the equation for each element of $\mathbf{r}$. This is given as

$$r_j = (\delta_0 H_j)\, s_j + \sum_{k=0, k\neq j}^{N_c-1} (\delta_{k-j} H_k)\, s_k + w_j, \tag{4.8}$$

where H_j are the diagonal elements of the diagonal matrix $\mathbf{H}$. As can be seen in (4.8), the desired symbol s_j is corrupted by the rotational component $\delta_0 H_j$ which is the CPE, and the added additive noise is represented by the second term in (4.8) which is the ICI.

The basic ideology in [77] and [111] is to treat the ICI as added receiver noise and to estimate δ_0 while assuming knowledge of H_j. In [77], a least-squares estimator using pilot symbols is used to arrive at an estimate of δ_0. The drawback of the least-squares approach is that it simply treats the ICI as added noise and does not utilize any a-priori information. The method works well only at high SNR regions, i.e., when the ICI power is

low. In [111], the MLE of δ_0 is derived where the assumption of ICI plus the receiver noise begin additive white Gaussian is used. Effectively, the MLE makes use of second-order information on the ICI term in the form of its variance.

The CPE-based approaches of [77] and [111] are attractive for the simplicity of implementing the estimators. The ease of implementation mainly comes from the fact that only one component of the vector $\boldsymbol{\delta}$ needs to be estimated. These methods, however, work well only for slow-varying phase noise processes.

Blind CPE estimation of [112]

A blind method for CPE estimation is proposed in [112], where the OFDM link is impaired by frequency offset. Such an approach can also be applied for the phase noise scenario, however, only for very low phase noise levels. The basic idea is to utilize the fact that the outermost points of an M-QAM constellation resemble that of a QPSK constellation. In Fig. 4.2, a scatter plot of the equalized symbols, i.e., $z_j = r_j/H_j$ is shown. Focusing on points outside the circle, we see that they resemble a rotated QPSK constellation. Taking the average of these outermost points will yield points on the rotated axis using which one can determine the CPE.

The advantage of such a method is that it does not require any knowledge of pilot symbols nor does it use a decision-directed approach. The disadvantage, however, is that it can only be applied for very small phase noise levels and works well to remove carrier frequency offset.

Power series-based estimation of [113]

In [113], the authors address the problem of phase noise estimation using a power series model for the phase noise process. Thus, the estimation step boils down to estimating the parameters of the power series model. The authors also propose a linear MMSE based channel estimator taking into account the effect of phase noise. This is discussed in Section 4.1.2.

Consider the OFDM system model given by (4.1) which can be expressed in terms of the channel frequency response H_j as

$$\mathbf{r}^m = \mathbf{S}^m\tilde{\mathbf{H}} + \mathbf{v}^m + \mathbf{w}, \tag{4.9}$$

where the superscript refers to the m-th OFDM symbol. The channel vector with CPE is $\tilde{\mathbf{H}} = \delta_0^m \, [H_0 \; H_1 \; \ldots \; H_{N_\mathrm{c}-1}]^\mathrm{T}$ and $\mathbf{S}^m$ is a diagonal matrix with elements s_j^m. The vector $\mathbf{v}^m$ denotes the ICI vector, and the goal here is to estimate this vector assuming an estimate of $\tilde{\mathbf{H}}$ and $\mathbf{S}^m$ is available.

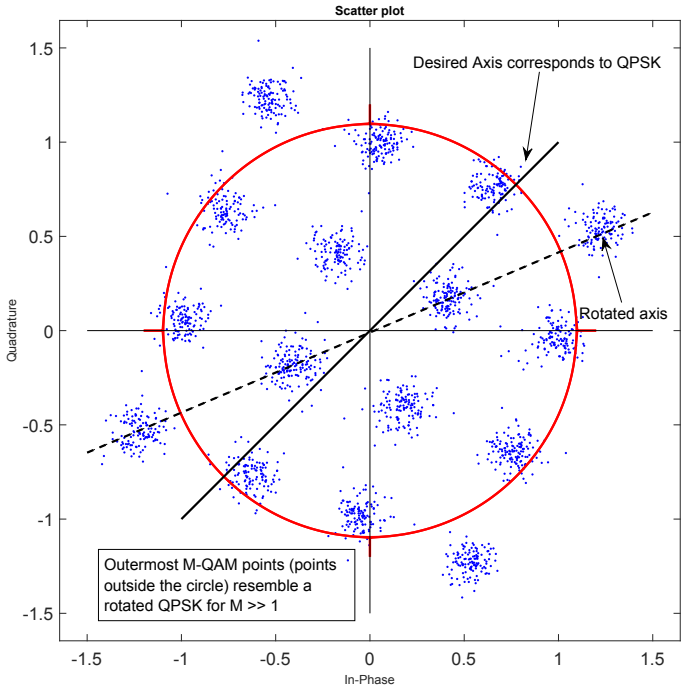


Figure 4.2. Scatter plot of the equalized symbols $z_j = r_j/H_j$. Symbol constellation used is 16-QAM. The Wiener phase noise 3-dB bandwidth is set to 100 Hz with $N_c = 2048$ and bandwidth equal to 30 MHz. The SNR is set to 30 dB.

This is done using a parametric model, where $\mathbf{v}^m$ is characterized using a certain number of parameters. The parametric model used is the power series model for the phase noise realization, $\theta^m[n]$, i.e.,

$$\theta^m[n] = \sum_{i=0}^{p} \lambda_i^m n^i, \tag{4.10}$$

where λ_i^m are the parameters for the m-th OFDM symbol. By using the small phase noise approximation, i.e., $e^{\jmath\theta^m[n]} \approx 1 + \jmath\theta^m[n]$ for $\theta^m[n] \approx 0$, the vector $\mathbf{v}^m$ can then be expressed in terms of λ_i^m as done in [113]. Finally, the ML estimate of these parameters is then derived.

Linear interpolation approach of [114]

Two simple yet extremely effective schemes for phase noise estimation are proposed in [114]. The motivation for these methods were mainly to improve the basis expansion approach of [110]. These methods, however, can be applied to improve any other phase noise estimation scheme that relies on *block-based* processing which is processing of OFDM symbols block-by-block. Phase noise processes are continuous in nature, and the

use of block-based processing for estimation of the phase noise realization results in discontinuities at the borders of consecutive OFDM symbols.

To improve phase noise estimates at symbol transitions, the method in [114] interpolates between phase noise estimates of consecutive OFDM symbols. Such an interpolator is given by [115]

$$\hat{\theta}^m[N_{\rm c}-W+n] = \hat{\theta}^m[N_{\rm c}-W] + n\frac{(\hat{\theta}^{m+1}[W]-\hat{\theta}^m[N_{\rm c}-W])}{2W}, \tag{4.11}$$

$$\hat{\theta}^{m+1}[n] = \hat{\theta}^m[N_{\rm c}-W] + (W+n)\frac{(\hat{\theta}^{m+1}[W]-\hat{\theta}^m[N_{\rm c}-W])}{2W}, \tag{4.12}$$

where W is the number of samples on each side of border, and $\hat{\theta}^m$, $\hat{\theta}^{m+1}$ are the phase noise estimates of the mth and $(m+1)$th OFDM symbol, respectively. In [114], these phase noise estimates are obtained using the basis expansion approach of [110]. In [114], this interpolation approach is given the name *linear interpolation tail estimation*. The second method, known by the name *linear interpolation CPE estimation*, interpolates between the CPE estimates of current and consecutive OFDM symbol to obtain phase noise estimates for the entire OFDM symbol length.

Low-pass filtering approach of [116] and [117]

Another effective phase noise estimation scheme based on *low pass filtering* is proposed in [116]. We briefly summarize the approach. Consider the time domain OFDM signal model of (4.1) which is given by

$$\begin{aligned}\mathbf{y} = \mathbf{F}^\dagger\mathbf{r} &= \mathrm{diag}(\boldsymbol{\phi})\mathbf{F}^\dagger\mathbf{H}\mathbf{s} + \mathbf{F}^\dagger\mathbf{w} \\ &= \mathrm{diag}\left(\mathbf{F}^\dagger\mathbf{H}\mathbf{s}\right)\boldsymbol{\phi} + \mathbf{F}^\dagger\mathbf{w} \\ &= \mathrm{diag}\left(\mathbf{x}\right)\boldsymbol{\phi} + \mathbf{n},\end{aligned} \tag{4.13}$$

where $\boldsymbol{\phi} = \left[e^{j\theta[0]}, e^{j\theta[1]}, \ldots, e^{j\theta[N_{\rm c}-1]}\right]^{\rm T}$, $\mathbf{x} = \mathbf{F}^\dagger\mathbf{H}\mathbf{s}$ and $\mathbf{n} = \mathbf{F}^\dagger\mathbf{w}$. Assuming knowledge of $\mathbf{H}$ and $\mathbf{s}$, we then perform

$$\Phi = \mathrm{diag}\left(\mathbf{x}^*\right)\mathbf{y} = \mathrm{diag}\left(|\mathbf{x}|^2\right)\boldsymbol{\phi} + \mathrm{diag}\left(\mathbf{x}^*\right)\mathbf{n}, \tag{4.14}$$

where $\mathbf{x}^*$ is the conjugate of $\mathbf{x}$, and $|\mathbf{x}|^2$ is the vector whose elements are the squared-magnitude values of the elements of $\mathbf{x}$. The factor $\mathrm{diag}\left(|\mathbf{x}|^2\right)$ can be interpreted as a scaling factor that gives a larger weight to high SNR phase noise samples and a lower weight to low SNR samples. The next master-stroke step is the low pass filtering of the vector Φ by recognizing that $\boldsymbol{\phi}$ is in general a low pass process, and its components always have unit magnitude. This entails the following operations to yield an estimate of $\theta[n]$

$$\hat{\theta}[n] = \arg\left(\mathrm{LP}\left\{\Phi[n]\right\}\right), \tag{4.15}$$

where $\Phi[n]$ denote the components of Φ, $\mathrm{LP}\{\cdot\}$ is the low-pass filtering operation.

4.1.2 Channel Estimation Including Phase Noise

In the previous section, we discussed phase noise estimation schemes to enable effective symbol estimation and detection. The estimation schemes, however, require knowledge of the channel. Initial studies on phase noise estimation for OFDM assumed that channel estimates could be obtained using traditional approaches, for example, interpolation between scattered pilots. These approaches do not take phase noise into account, and channel estimates obtained using these approaches will deteriorate as the levels of phase noise increase. This observation has led to works dedicated to channel estimation in the presence of phase noise. Let us now review some of these works.

The MAP-based joint estimator of [118]

In [118], using a preamble OFDM symbol, a joint channel, phase noise and frequency offset estimator based on the maximum a-posteriori probability (MAP) criteria is proposed. Consider the time-domain OFDM signal model which is given by

$$\mathbf{y} = \mathrm{diag}\left(\boldsymbol{\varepsilon}\right)\mathrm{diag}(\boldsymbol{\phi})\mathbf{F}^{\dagger}\mathbf{S}\mathbf{F}_{\mathrm{t}}\mathbf{h} + \mathbf{n}, \tag{4.16}$$

where $\boldsymbol{\varepsilon} = \left[1, e^{\jmath 2\pi\epsilon/N_{\mathrm{c}}}, \ldots, e^{\jmath 2\pi(N_{\mathrm{c}}-1)\epsilon/N_{\mathrm{c}}}\right]^{\mathrm{T}}$ with ϵ denoting the normalized carrier frequency offset, $\mathbf{h} = [h[0],\ h[1], \ldots, h[L-1]]^{\mathrm{T}}$ whose elements are the IDFT of H_j, $\boldsymbol{\phi} = \left[e^{j\theta[0]},\ e^{j\theta[1]}, \ldots, e^{j\theta[N_{\mathrm{c}}-1]}\right]^{\mathrm{T}}$, $\mathbf{S} = \mathrm{diag}\left(\mathbf{s}\right)$, and finally $\mathbf{n}$ denotes the receiver white noise. The MAP estimator seeks those parameter values that maximizes the a-posterior probability [104], i.e.,

$$\left(\hat{\mathbf{h}}, \hat{\boldsymbol{\theta}}, \hat{\epsilon}\right) = \arg\max_{\mathbf{h},\boldsymbol{\theta},\epsilon} p\left(\mathbf{h}, \boldsymbol{\theta}, \epsilon|\mathbf{y}\right), \tag{4.17}$$

where $\boldsymbol{\theta} = [\theta[0],\ \theta[1], \ldots, \theta[N_{\mathrm{c}}-1]]^{\mathrm{T}}$. Maximizing the above function is equivalent to minimizing the *negative log-likelihood* which is given by

$$\mathcal{L}\left(\mathbf{h}, \boldsymbol{\theta}, \epsilon\right) = -\log p\left(\mathbf{y}|\mathbf{h}, \boldsymbol{\theta}, \epsilon\right) - \log p\left(\boldsymbol{\theta}\right), \tag{4.18}$$

where we used the Bayes rule $p(x|y) \propto p(y|x)p(x)$ in the above equation, and we treat $\mathbf{h}$ and ϵ as deterministic quantities, while $\boldsymbol{\theta}$ is assumed to be drawn from a prior distribution that is usually Gaussian. Using (4.16), it can be easily seen that the conditional distribution in (4.18) is also Gaussian. Deriving closed-form analytical expressions for the estimators by

minimizing (4.18) is perhaps analytically intractable. Suboptimal estimates can be obtained by minimizing $\mathcal{L}$ w.r.t. a certain parameter while fixing other parameters and then performing *forward substitution* to obtain a reduced likelihood function that depends only on the fixed parameters. The process is then repeated until we have a likelihood function only in one variable. We now show that such an approach always yields suboptimal estimates.

Let $\mathbf{h}^\star, \epsilon^\star$ and $\boldsymbol{\theta}^\star$ denote the global minimizers to $\mathcal{L}$. We then have that

$$\mathcal{L}\left(\mathbf{h}^\star, \boldsymbol{\theta}^\star, \epsilon^\star\right) \leq \mathcal{L}\left(\mathbf{h}, \boldsymbol{\theta}, \epsilon\right). \tag{4.19}$$

Fixing ϵ and $\boldsymbol{\theta}$, we minimize $\mathcal{L}$ w.r.t. $\mathbf{h}$. Let $\mathbf{h}^\diamond$ be the minimizer. We then must have

$$\mathcal{L}\left(\mathbf{h}^\star, \boldsymbol{\theta}^\star, \epsilon^\star, \right) \leq \mathcal{L}\left(\mathbf{h}^\diamond, \boldsymbol{\theta}, \epsilon\right) \leq \mathcal{L}\left(\mathbf{h}, \boldsymbol{\theta}, \epsilon\right). \tag{4.20}$$

Let $\boldsymbol{\theta}^\diamond$ be the minimizer to $\mathcal{L}\left(\mathbf{h}^\diamond, \boldsymbol{\theta}, \epsilon\right)$, where the minimization is done w.r.t. $\boldsymbol{\theta}$ while keeping ϵ fixed which again leads to

$$\mathcal{L}\left(\mathbf{h}^\star, \boldsymbol{\theta}^\star, \epsilon^\star\right) \leq \mathcal{L}\left(\mathbf{h}^\diamond, \boldsymbol{\theta}^\diamond, \epsilon\right) \leq \mathcal{L}\left(\mathbf{h}^\diamond, \boldsymbol{\theta}, \epsilon\right) \leq \mathcal{L}\left(\mathbf{h}, \boldsymbol{\theta}, \epsilon\right). \tag{4.21}$$

Let $\epsilon^\diamond$ be the minimizer to $\mathcal{L}\left(\mathbf{h}^\diamond, \boldsymbol{\theta}^\diamond, \epsilon\right)$ which finally leads to the inequality

$$\mathcal{L}\left(\mathbf{h}^\star, \boldsymbol{\theta}^\star, \epsilon^\star\right) \leq \mathcal{L}\left(\mathbf{h}^\diamond, \boldsymbol{\theta}^\diamond, \epsilon^\diamond\right). \tag{4.22}$$

Closed-form expressions for $\mathcal{L}\left(\mathbf{h}^\diamond, \boldsymbol{\theta}, \epsilon\right)$ and $\mathcal{L}\left(\mathbf{h}^\diamond, \boldsymbol{\theta}^\diamond, \epsilon\right)$ and the estimators $\mathbf{h}^\diamond$ and $\boldsymbol{\theta}^\diamond$ are derived in [118]. The estimator $\epsilon^\diamond$ is obtained numerically using an exhaustive search over $\mathcal{L}\left(\mathbf{h}^\diamond, \boldsymbol{\theta}^\diamond, \epsilon\right)$.

The ML-based joint estimator of [119]

In [119], the authors utilize the same approach as [118], i.e., forward and backward substitution. There are, however, distinct differences. Firstly, the estimation is performed in the frequency domain, and joint ML estimates are derived unlike in [118] which seeks MAP estimates. Thus, no a-priori information on the phase noise process is used in [119]. Rather than estimating the phase noise process itself, the spectral components of the complex exponential of the phase noise process plus frequency offset are estimated. Hence, separate frequency offset and phase noise estimation steps are not needed. What now follows is a brief summary of the approach.

Consider the frequency-domain OFDM system model impaired by receiver phase noise:

$$\mathbf{r} = \mathbf{V}\mathbf{S}\mathbf{F}_{\mathrm{t}}\mathbf{h} + \mathbf{w}, \tag{4.23}$$

where $\mathbf{S}$ is diagonal with preamble s_j, and $\mathbf{h} = [h[0],\ h[1], \ldots, h[L-1]]^{\mathrm{T}}$ is the channel impulse response vector. The unknowns in the above equation are $\mathbf{h}$ and $\boldsymbol{\delta}$ which is used to form the unitary circulant $\mathbf{V}$ matrix. The ML estimate is obtained by minimizing the negative log-likelihood as follows

$$\mathcal{L}(\mathbf{h}, \boldsymbol{\delta}) = \arg \min_{\boldsymbol{\delta}, \mathbf{h}} \log p(\mathbf{r}|\mathbf{h}, \boldsymbol{\delta}). \tag{4.24}$$

As done in [118], a channel estimate is obtained by minimizing $\mathcal{L}(\mathbf{h}, \boldsymbol{\delta})$ only w.r.t. $\mathbf{h}$. Using the fact that $p(\mathbf{r}|\mathbf{h}, \boldsymbol{\delta})$ is Gaussian, the channel estimate can be easily derived to obtain

$$\hat{\mathbf{h}}_{\mathrm{ML}} = \left(\mathbf{F}_{\mathrm{t}}^{\dagger}\mathbf{S}^{\dagger}\mathbf{S}\mathbf{F}_{\mathrm{t}}\right)^{-1}\mathbf{F}_{\mathrm{t}}^{\dagger}\mathbf{S}^{\dagger}\mathbf{V}^{\dagger}\mathbf{r}. \tag{4.25}$$

By substituting (4.25) back into $\mathcal{L}(\mathbf{h}, \boldsymbol{\delta})$ and utilizing $\mathbf{V}^{\dagger}\mathbf{V} = \mathbf{I}_{N_{\mathrm{c}}}$, we obtain the likelihood function in terms of $\boldsymbol{\delta}$ which is given by

$$\mathcal{L}(\boldsymbol{\delta}) = \boldsymbol{\delta}^{\dagger}\mathbf{M}\boldsymbol{\delta}, \tag{4.26}$$

where $\mathbf{M} = \left(\mathbf{R}^{\dagger}\mathbf{R} - \mathbf{R}^{\dagger}\mathbf{P}_{\mathrm{r}}\mathbf{R}\right)^{\mathrm{T}}$ is Hermitian and $\mathbf{R}$ is a column-wise circulant matrix with its first column vector being $\mathbf{r}$. The orthogonal projection matrix is given by $\mathbf{P}_{\mathrm{r}} = \mathbf{S}\mathbf{F}_{\mathrm{t}}\mathbf{B}^{-1}\mathbf{F}_{\mathrm{t}}^{\dagger}\mathbf{S}^{\dagger}$ with $\mathbf{B} = \mathbf{F}_{\mathrm{t}}^{\dagger}\mathbf{S}^{\dagger}\mathbf{S}\mathbf{F}_{\mathrm{t}}$.

From (4.25), we see that the channel estimate requires knowledge of the matrix $\mathbf{V}$. This is obtained by minimizing the quadratic likelihood function $\mathcal{L}(\boldsymbol{\delta})$ whose estimate is then used to form an estimate of $\mathbf{V}$. We now focus our attention on minimizing $\mathcal{L}(\boldsymbol{\delta})$. Since $\mathcal{L}(\boldsymbol{\delta})$ is a homogeneous quadratic cost function, the minimizer is the trivial null vector of zeros. In order to obtain a sensible estimate, a constraint needs to be enforced. In [119], the authors propose to use a linear constraint. Specifically, the minimization problem is given by

$$\text{Minimize } \mathcal{L}(\boldsymbol{\delta}) = \boldsymbol{\delta}^{\dagger}\mathbf{M}\boldsymbol{\delta} \tag{4.27}$$

$$\text{such that } \frac{1}{2}\left(\boldsymbol{\delta}^{\dagger}\mathbf{e}_1 + \mathbf{e}_1^{\dagger}\boldsymbol{\delta}\right) = 1, \tag{4.28}$$

where the $N_{\mathrm{c}} \times 1$ vector $\mathbf{e}_1 = [1, 0, \ldots, 0]^{\mathrm{T}}$. The constraint in (4.28) requires that the minimizer to $\mathcal{L}(\boldsymbol{\delta})$ has maximum correlation with $\mathbf{e}_1$. Ideally, when no phase noise is present, we have $\boldsymbol{\delta} = \mathbf{e}_1$. In practice, phase noise is always present, however, if the phase noise process is slowly varying then there is strong correlation between $\boldsymbol{\delta}$ and $\mathbf{e}_1$, and, thus, the constraint in (4.28) is applicable.

The MMSE channel estimator of [113]

In [113], in addition to phase noise estimation, the authors also address the problem of channel estimation in the presence of phase noise. By

treating the ICI term as added receiver noise with known second-order statistics, a linear MMSE channel estimator is derived. Let us briefly describe the method.

Equation (4.1) can be expressed in terms of H_j as

$$\mathbf{r}^m = \mathbf{S}^m \tilde{\mathbf{H}} + \mathbf{v}^m + \mathbf{w}, \tag{4.29}$$

where $\tilde{\mathbf{H}} = \delta_0^m \left[H_0 \; H_1 \; \ldots \; H_{N_c-1}\right]^{\mathrm{T}}$, $\mathbf{S}^m$ is diagonal with elements s_j^m, $\mathbf{v}^m$ is the ICI vector, and the superscript refers to the mth OFDM symbol. If the channel is quasi-static, i.e., it does not change for a few OFDM symbols then previous OFDM symbols can also be used in the estimation of $\tilde{\mathbf{H}}$ as follows: Using previous symbol decisions, $\hat{\mathbf{S}}^{m-k}$, form the vector

$$\mathbf{y} = \left[\left(\mathbf{y}^m\right)^{\mathrm{T}}, \; \left(\mathbf{y}^{m-1}\right)^{\mathrm{T}}, \ldots, \left(\mathbf{y}^{m-K}\right)^{\mathrm{T}}\right]^{\mathrm{T}} \tag{4.30}$$

where $\mathbf{y}^{m-k} = \left(\hat{\mathbf{S}}^{m-k}\right)^{-1} \mathbf{r}^{m-k}, \; k = 1, 2, \ldots, K$. For the mth OFDM symbol, $\mathbf{y}^m = \left(\mathbf{S}_{\mathrm{p}}^m\right)^{-1} \mathbf{r}_{\mathrm{p}}^m$ where $\mathbf{S}_{\mathrm{p}}^m$ is diagonal comprising of the pilot symbols for the mth OFDM symbol, and $\mathbf{r}_{\mathrm{p}}^m$ is the received signal vector corresponding to the pilot symbols. The linear MMSE estimate of $\tilde{\mathbf{H}}$ is

$$\hat{\tilde{\mathbf{H}}} = \mathbf{C}_{\tilde{\mathbf{H}}\mathbf{y}} \mathbf{C}_{\mathbf{y}\mathbf{y}}^{-1} \mathbf{y}. \tag{4.31}$$

Closed-form expressions for the cross-covariance and covariance matrices in the above equation are derived and can be found in [113].

Monte Carlo-based EM channel estimation of [120] and [121]

In [120] and [121], the expectation maximization (EM) algorithm is used to estimate the channel in the presence of phase noise and carrier frequency offset. In [120], the estimation is performed for a generic multicarrier system while, in [121], the results are derived for OFDM [122]. In addition to estimating the channel, the work in [121] estimates the receiver noise variance, while the work in [120] goes a step further to also estimate the 3-dB bandwidth of phase noise. These parameters are estimated as part of the EM framework. To implement the EM-algorithm, *Monte Carlo* methods are used in [120] and [121] which are numerical methods for obtaining probability density functions. We now summarize the ideology of [120] and [121] using OFDM as an example.

Consider the time-domain OFDM system model impaired by phase noise and frequency offset:

$$\mathbf{y} = \mathrm{diag}\left(\boldsymbol{\varepsilon}\right) \mathrm{diag}(\boldsymbol{\phi}) \mathbf{G} \left[\mathbf{I}_{N_c} \; \mathbf{I}_{N_c}\right] \begin{pmatrix} \mathbf{x}_{\mathrm{d}} \\ \mathbf{x}_{\mathrm{p}} \end{pmatrix} + \mathbf{n}, \tag{4.32}$$

where $\boldsymbol{\varepsilon} = \left[1, e^{\jmath 2\pi\epsilon/N_\mathrm{c}}, \ldots, e^{\jmath 2\pi(N_\mathrm{c}-1)\epsilon/N_\mathrm{c}}\right]^\mathrm{T}$ with ϵ denoting the normalized carrier frequency offset, and $\boldsymbol{\phi} = \left[e^{\jmath\theta[0]},\ e^{\jmath\theta[1]}, \ldots, e^{\jmath\theta[N_\mathrm{c}-1]}\right]^\mathrm{T}$. The channel matrix $\mathbf{G}$ is circulant with $\mathbf{h} = [h[0],\ h[1], \ldots, h[L-1]]^\mathrm{T}$. The vectors $\mathbf{x}_\mathrm{d} = \mathbf{F}^\dagger \mathbf{s}_\mathrm{d}$ and $\mathbf{x}_\mathrm{p} = \mathbf{F}^\dagger \mathbf{s}_\mathrm{p}$ represent the data portions and pilot portions of the OFDM signal, where the transmitted symbol vector $\mathbf{s} = \mathbf{s}_\mathrm{d} + \mathbf{s}_\mathrm{p}$. The vector $\mathbf{n}$ is white Gaussian with variance σ_n^2.

In [120], a Wiener process is assumed for $\theta[n]$, i.e.,

$$\theta[n] = \theta[n-1] + \nu[n], \tag{4.33}$$

where $\nu[n] \sim \mathcal{N}\left(0, 2\pi\frac{f_{3\mathrm{dB}}}{f_\mathrm{sub}N_\mathrm{c}}\right)$ with $f_{3\mathrm{dB}}$ and f_sub denoting, respectively, the phase noise 3-dB bandwidth and OFDM subcarrier spacing. In most works, it is assumed that $f_{3\mathrm{dB}}$ is known while, in [120], this parameter is assumed unknown and estimated.

Denote the parameter vector to be estimated by $\Gamma = \left[\mathbf{h}^\mathrm{T}, \epsilon, f_{3\mathrm{dB}}, \sigma_\mathrm{n}^2\right]^\mathrm{T}$. These parameters are estimated using the EM algorithm which is an iterative method of obtaining ML estimates in the presence of *hidden variables* [123]. In this case, the vector $\mathbf{z} = \left[\mathbf{x}_\mathrm{d}^\mathrm{T},\ \boldsymbol{\theta}^\mathrm{T}\right]^\mathrm{T}$ is treated as the hidden variable, where $\boldsymbol{\theta} = [\theta[0],\ \theta[1], \ldots, \theta[N_\mathrm{c}-1]]^\mathrm{T}$. Random parameters that are not directly observable are generally treated as hidden variables. The EM algorithm iterates between an *expectation step (E-step)* and a *maximization step (M-step)*. Specifically, these steps are

$$\text{E-step:}\quad \mathcal{L}\left(\Gamma, \Gamma_i\right) = \int p\left(\mathbf{z}|\mathbf{y}, \Gamma_i\right) \log p\left(\mathbf{z}, \mathbf{y}|\Gamma\right) \mathrm{d}\mathbf{z} \tag{4.34}$$

$$\text{M-step:}\qquad \Gamma_{i+1} = \arg\max_{\Gamma} \mathcal{L}\left(\Gamma, \Gamma_i\right). \tag{4.35}$$

In general, closed-form expressions for the E-step and M-step are not easily available, where, typically, the difficulty is the non-Gaussianity of the posterior PDF of the hidden variable $\mathbf{z}$, i.e., $p\left(\mathbf{z}|\mathbf{y}, \Gamma_i\right)$. Expressing the PDF in terms of $\mathbf{x}_\mathrm{d}$ and $\boldsymbol{\theta}$, we have

$$p\left(\mathbf{z}|\mathbf{y}, \Gamma_i\right) \propto p\left(\mathbf{x}_\mathrm{d}|\mathbf{y}, \boldsymbol{\theta}, \Gamma_i\right) p\left(\boldsymbol{\theta}|\mathbf{y}, \Gamma_i\right) \tag{4.36}$$

$$\propto p\left(\mathbf{y}|\mathbf{x}_\mathrm{d}, \boldsymbol{\theta}, \Gamma_i\right) p\left(\mathbf{x}_\mathrm{d}\right) p\left(\boldsymbol{\theta}|\mathbf{y}, \Gamma_i\right), \tag{4.37}$$

where we have used the fact that $\mathbf{x}_\mathrm{d}$ and $\boldsymbol{\theta}$ are independent of each other. The difficulty is because of non-Gaussianity of $p\left(\boldsymbol{\theta}|\mathbf{y}, \Gamma_i\right) \propto p\left(\mathbf{y}|\boldsymbol{\theta}, \Gamma_i\right) p\left(\boldsymbol{\theta}\right)$ which is due to the non-linear relationship between $\mathbf{y}$ and $\boldsymbol{\theta}$, where the non-linear function is the complex exponential function. Since $p\left(\mathbf{x}_\mathrm{d}\right)$ and $p\left(\mathbf{y}|\mathbf{x}_\mathrm{d}, \boldsymbol{\theta}, \Gamma_i\right)$ are Gaussian, and $\mathbf{y}$ is linear in $\mathbf{x}_\mathrm{d}$, their product can be represented using a Gaussian distribution.

Typically, direct computation of $p(\mathbf{z}|\mathbf{y}, \Gamma_i)$ can be challenging from the point of view of computational complexity because the dimensionality of the hidden variables $\mathbf{z}$ and observation vector $\mathbf{y}$ can be large. Less complex approaches are recursive methods of obtaining the joint PDF; this is neatly summarized in the following equation [104]

$$p(z_{0:k}|y_{0:k}) \propto p(y_k|z_{0:k}, y_{0:k-1})\, p(z_{0:k}|y_{0:k-1}) \tag{4.38}$$

$$\propto p(y_k|z_k)\, p(z_k|z_{0:k-1}, y_{0:k-1})\, p(z_{0:k-1}|y_{0:k-1}) \tag{4.39}$$

$$\propto p(y_k|z_k)\, p(z_k|z_{0:k-1})\, p(z_{0:k-1}|y_{0:k-1}), \tag{4.40}$$

where $x_{0:k}$ denote elements $x_0, x_1, \ldots, x_k$. In arriving at (4.40), we made use of the so-called *Markov property* [104]. From (4.40), we see that the joint posterior-i PDF at iteration k can be obtained from joint posteriori PDF at iteration $k-1$ assuming a certain measurement model represented by $p(y_k|z_k)$ and a state-space model represented by $p(z_k|z_{0:k-1})$.

Similar to the direct approach, the recursive computation is hindered by the non-Gaussianity of $p(\theta_{0:k}|y_{0:k})$. In cases where the posteriori PDF is Gaussian, Kalman filters can be applied to recursively obtain the posteriori PDF. In non-Gaussian cases, linearization techniques such as the extended Kalman filter and statistically linearized filters can be applied [104]. The last resort is to numerically evaluate $p(\theta_{0:k}|y_{0:k})$ by drawing samples that are representative of the distribution. Such methods are referred to as Monte Carlo methods with *particle filtering* being a popular technique. An excellent, precise and concise treatment of this subject can be found in [104, 120]. To summarize, the work in [120], computes EM-estimates using (4.34), where the posteriori PDFs are computed using recursive Monte Carlo methods.

The EM-based joint estimator of [124]

Another EM-based joint estimation of channel, phase noise and frequency offset is proposed in [124]. It mainly aims to address the shortcomings of [118] which is its high computational complexity. Rather than obtain MAP estimates, the goal here is, using the EM algorithm, to obtain ML estimates of channel and frequency offset while treating phase noise as the hidden variable. We briefly summarize the approach.

The time-domain OFDM system model in consideration is given by

$$\mathbf{y} = \operatorname{diag}(\boldsymbol{\varepsilon}) \operatorname{diag}(\boldsymbol{\phi}) \mathbf{F}^{\dagger} \mathbf{S} \mathbf{F}_{\mathrm{t}} \mathbf{h} + \mathbf{n}, \tag{4.41}$$

where $\boldsymbol{\varepsilon} = \left[1, e^{\jmath 2\pi\epsilon/N_{\mathrm{c}}}, \ldots, e^{\jmath 2\pi(N_{\mathrm{c}}-1)\epsilon/N_{\mathrm{c}}}\right]^{\mathrm{T}}$ with ϵ denoting the normalized carrier frequency offset, $\mathbf{h} = [h[0],\ h[1], \ldots, h[L-1]]^{\mathrm{T}}$ whose elements are

the IDFT of H_j, $\boldsymbol{\phi} = \left[e^{j\theta[0]},\ e^{j\theta[1]}, \ldots, e^{j\theta[N_c-1]}\right]^{\mathrm{T}}$, $\mathbf{S} = \mathrm{diag}\,(\mathbf{s})$, and finally $\mathbf{n}$ denotes the receiver white noise. Let $\Gamma = \left[\mathbf{h}^{\mathrm{T}}, \epsilon\right]^{\mathrm{T}}$ denote the parameter vector to be estimated and $\boldsymbol{\theta} = [\theta[0],\ \theta[1], \ldots, \theta[N_c - 1]]^{\mathrm{T}}$ denotes the hidden variable. The E-step and M-step are given as follows

$$\text{E-step:} \quad \mathcal{L}\left(\Gamma, \Gamma_i\right) = \int p\left(\boldsymbol{\theta}|\mathbf{y}, \Gamma_i\right) \log p\left(\boldsymbol{\theta}, \mathbf{y}|\Gamma\right) \mathrm{d}\boldsymbol{\theta} \tag{4.42}$$

$$\text{M-step:} \quad \Gamma_{i+1} = \arg\max_{\Gamma} \mathcal{L}\left(\Gamma, \Gamma_i\right). \tag{4.43}$$

Ideally, with complete knowledge of $\boldsymbol{\theta}$, the term $\log p\left(\boldsymbol{\theta}, \mathbf{y}|\Gamma\right)$ in (4.42) represents the likelihood function. However, we do not know $\boldsymbol{\theta}$ and obtain knowledge of it using the a-posteri PDF given by $p\left(\boldsymbol{\theta}|\mathbf{y}, \Gamma_i\right)$ in (4.42). This a-posteri PDF of $\boldsymbol{\theta}$ is obtained using the extended Kalman filter [104]. The usage of a Kalman filter implicitly assumes a Gaussian a-posteri PDF. In [124], the authors obtain hard estimate of $\boldsymbol{\theta}$ using this a-posteriori PDF, i.e.,

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} p\left(\boldsymbol{\theta}|\mathbf{y}, \Gamma_i\right). \tag{4.44}$$

The above estimator is a MAP estimator and, for a Gaussian distribution, it is also the MMSE estimator. With this estimate of $\boldsymbol{\theta}$ available, the cost function to be maximized is modified to

$$\tilde{\mathcal{L}}\left(\Gamma, \Gamma_i\right) = \log p\left(\hat{\boldsymbol{\theta}}, \mathbf{y}|\Gamma\right). \tag{4.45}$$

Comparing the above equation with (4.42), we see that the averaging operation in (4.42) is replaced by a point-density estimate which is the MAP or MMSE estimate of the a-posteriori Gaussian density.

4.1.3 Joint Phase Noise and Symbol Estimation

In this section, we review some state-of-the-art methods for joint phase noise and symbol estimation during the data phase of transmission. In Section 4.1.1, we reviewed some estimation schemes where phase noise estimation and symbol estimation/detection are performed independently of each other. This implies that such approaches do not necessarily deliver *statistically optimal* phase noise and symbol estimates. Statistically optimal estimates can be obtained by performing simultaneous phase noise and symbol estimation using Bayesian inference methods.

The variational inference approach of [125]

In [125], the authors address the problem of joint symbol and phase noise estimation using *variational inference*. It is an approximation to Bayesian

inference, and a good overview of the subject can be found in [126, 127]. We now briefly summarize the approach of [125].

Consider the received time-domain OFDM signal in the presence of phase noise which is given by

$$\mathbf{y} = \operatorname{diag}\left(\mathbf{F}^{\dagger}\mathbf{H}\mathbf{s}\right)\boldsymbol{\phi} + \mathbf{n} \tag{4.46}$$

$$\approx \operatorname{diag}\left(\mathbf{F}^{\dagger}\mathbf{H}\mathbf{s}\right)(\mathbf{1} + \jmath\boldsymbol{\theta}) + \mathbf{n}, \tag{4.47}$$

where $\boldsymbol{\phi} = \left[e^{\jmath\theta[0]},\ e^{\jmath\theta[1]}, \ldots, e^{\jmath\theta[N_{\mathrm{c}}-1]}\right]^{\mathrm{T}}$ and $\boldsymbol{\theta} = [\theta[0],\ \theta[1], \ldots, \theta[N_{\mathrm{c}}-1]]^{\mathrm{T}}$. In the second step, we have used the small angle approximation, i.e., $e^{\jmath\theta} \approx (1 + \jmath\theta)$ which is reasonable for a slow varying phase noise process.

The statistically optimal estimates are obtained from the a-posteriori density of s which is given by

$$p\left(\mathbf{s}|\mathbf{y}\right) = \int p\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right) d\boldsymbol{\theta} \tag{4.48}$$

$$\propto \int p\left(\mathbf{y}|\mathbf{s}, \boldsymbol{\theta}\right) p\left(\mathbf{s}\right) p\left(\boldsymbol{\theta}\right) d\boldsymbol{\theta} \tag{4.49}$$

$$= p\left(\mathbf{s}\right) \int p\left(\mathbf{y}|\mathbf{s}, \boldsymbol{\theta}\right) p\left(\boldsymbol{\theta}\right) d\boldsymbol{\theta}, \tag{4.50}$$

where $p\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right)$ is the joint a-posteriori density of s and $\boldsymbol{\theta}$, and in arriving at (4.50), we use the fact that s and $\boldsymbol{\theta}$ are independent of each other.

The are two major difficulties in the evaluation of $p\left(\mathbf{s}|\mathbf{y}\right)$. The first difficulty has to do with $p\left(\mathbf{s}\right)$ which is discrete in nature because of the discrete nature of s. For example, each component of s is typically drawn from an M-QAM constellation. Thus, to obtain the MAP-estimate of s, we need to compare $M^{N_{\mathrm{c}}}$ values of $p\left(\mathbf{s}|\mathbf{y}\right)$ which clearly has exponential complexity. To reduce the complexity, $p\left(\mathbf{s}\right)$ is generally assumed to be drawn from some continuous distribution. The second difficulty is in the evaluation of the integral in (4.50). A resulting closed-form expression, assuming a Gaussian density for $\boldsymbol{\theta}$, can be obtained. However, this expression in terms of the variable s is highly complicated and its manipulation, for example to obtain MAP or MMSE estimates, is mathematically intractable.

To alleviate the aforementioned problems, an approximation to $p\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right)$ is sought such that the resulting expression can be easily manipulated. Denote $q\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right)$ to be our approximation, where the function q is chosen from a particular family. Furthermore, if we assume that $q\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right) = q\left(\mathbf{s}|\mathbf{y}\right) q\left(\boldsymbol{\theta}|\mathbf{y}\right)$, we readily have an approximation to $p\left(\mathbf{s}|\mathbf{y}\right)$. Thus, the following question is which members from the family, with the chosen factorization, yield the best approximation to $p\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right)$. Such an approach of approximation is known as variational inference which is used in [125].

We briefly summarize the ideology. Consider the log-likelihood function of (4.48) which is given by

$$\mathcal{L}(\mathbf{s}) = \log \int p\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right) d\boldsymbol{\theta} \tag{4.51}$$

$$\propto \log \int q\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right) \frac{p\left(\mathbf{s}, \boldsymbol{\theta}, \mathbf{y}\right)}{q\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right)} d\boldsymbol{\theta} \tag{4.52}$$

$$\geq \int q\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right) \log \frac{p\left(\mathbf{s}, \boldsymbol{\theta}, \mathbf{y}\right)}{q\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right)} d\boldsymbol{\theta} \tag{4.53}$$

$$\geq - \int q\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right) \log \frac{q\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right)}{p\left(\mathbf{s}, \boldsymbol{\theta}, \mathbf{y}\right)} d\boldsymbol{\theta}, \tag{4.54}$$

where we used that $p\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right) \propto p\left(\mathbf{s}, \boldsymbol{\theta}, \mathbf{y}\right)$. In the above equations, the distribution $q\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right)$ represents our approximation of $p\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right)$. In the last step, we have made use of the fact that the logarithm is a concave function. The MAP estimate is obtained by maximizing $\mathcal{L}(\mathbf{s})$ and, after observing (4.54), a suboptimal estimate is obtained by first maximizing the lower bound w.r.t. q. The optimal function $q\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right)$ that maximizes the lower bound is indeed $q^{\star} = p\left(\mathbf{s}, \theta|\mathbf{y}\right)$. However, as stated earlier, the distribution $p\left(\mathbf{s}, \theta|\mathbf{y}\right)$ poses various difficulties. By restricting $q\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right)$ to a certain family of distributions, we seek among that family the maximizer to the lower bound in (4.54).

Maximizing the lower bound in (4.54) w.r.t. q is equivalent to minimizing

$$\mathcal{K}\left(q\right) = \int_{\mathbf{s},\boldsymbol{\theta}} q\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right) \log \frac{q\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right)}{p\left(\mathbf{s}, \boldsymbol{\theta}, \mathbf{y}\right)} d\mathbf{s} d\boldsymbol{\theta}, \tag{4.55}$$

which is the *Kullback-Leibler* divergence between $q\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right)$ and $p\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right)$. Assuming $q\left(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y}\right) = q\left(\mathbf{s}|\mathbf{y}\right) q\left(\boldsymbol{\theta}|\mathbf{y}\right)$, we can now obtain these distributions by minimizing (4.55). As an example, assume that q comes from the Gaussian family, i.e.,

$$q\left(\mathbf{s}|\mathbf{y}\right) = \mathcal{CN}\left(\mathbf{m}_{\mathbf{s}}, \mathbf{C}_{\mathbf{s}}\right), \tag{4.56}$$

$$q\left(\boldsymbol{\theta}|\mathbf{y}\right) = \mathcal{N}\left(\mathbf{m}_{\boldsymbol{\theta}}, \mathbf{C}_{\boldsymbol{\theta}}\right), \tag{4.57}$$

where $\mathcal{CN}$ denotes complex Gaussian distribution and $\mathbf{m}_{\mathbf{s}}$, $\mathbf{m}_{\boldsymbol{\theta}}$, $\mathbf{C}_{\mathbf{s}}$ and $\mathbf{C}_{\boldsymbol{\theta}}$ denote the parameters of the distributions that need to optimized. Minimizing (4.55) w.r.t. q essentially transforms to minimizing w.r.t. the parameters in (4.56) and (4.57). In [125], closed-form expressions for $\mathbf{m}_{\mathbf{s}}, \mathbf{m}_{\boldsymbol{\theta}}, \mathbf{C}_{\mathbf{s}}$ and $\mathbf{C}_{\boldsymbol{\theta}}$ are derived by minimizing $\mathcal{K}\left(q\right)$ w.r.t. these parameters. The MAP estimates can then be derived from the obtained distributions. The Gaussian family is of course the easiest choice to work with. The authors also consider point densities which essentially lead to obtaining hard estimates.

4.2 Contributions to Phase Noise Estimation in OFDM

In this thesis, two novel phase noise estimation schemes are developed in Publications V and VI. The work in Publication VI falls in the category of joint channel and phase noise estimation, where the goal is to obtain accurate channel estimates in the presence of phase noise. It is an improvement to the original method proposed in [119] which is summarized in Section 4.1.2 under the title "The ML-based joint estimator of [119]". This improvement is achieved through the use of a *subspace-based* approach, where possible subspaces in which the phase noise spectral vector may lie are exploited during phase noise estimation.

The second phase noise estimation method, proposed in Publication V, can be viewed as an isolated approach, where the goal is to obtain accurate phase noise estimates after which phase noise is removed, and then symbol estimation/detection is performed. The work builds upon the work of [119], where a phase noise estimate is obtained by minimizing a homogeneous quadratic cost function subject to linear constraints. In Publication V, a *geometry-based* approach is used, where it is shown that the phase noise spectral vector always adheres to a specific type of geometry which mathematically is described by a set of *non-convex* quadratic equations. The work demonstrates that better phase noise estimates can be obtained by restricting the search space to this non-convex set rather than the convex set described by the linear constraints. The geometry-based method of Publication V uses a decision-directed approach. This research work is continued in [128], where the phase noise geometry is used in developing pilot-based approaches which generally are advantageous of low computational complexity and, hence, low latency.

The following sections summarize the ideology of the works of Publications V and VI. We begin with the subspace-based approach.

4.3 Subspace-based Phase Noise Estimation of Publication VI

In [119], a ML-based joint channel and phase noise estimator using a preamble symbol is derived, see Section 4.1.2 under the title "The ML-based joint estimator of [119]" for details. The channel and phase noise estimators are given by

$$\hat{\mathbf{h}}_{\mathrm{ML}} = \left(\mathbf{F}_{\mathrm{t}}^{\dagger}\mathbf{S}^{\dagger}\mathbf{S}\mathbf{F}_{\mathrm{t}}\right)^{-1}\mathbf{F}_{\mathrm{t}}^{\dagger}\mathbf{S}^{\dagger}\mathbf{V}^{\dagger}\mathbf{r}, \tag{4.58}$$

where the matrix $\mathbf{V}$ is unitary circulant with the first column vector $\boldsymbol{\delta}$. An estimate of $\boldsymbol{\delta}$ is obtained by solving the following optimization problem:

$$\text{Minimize } \mathcal{L}(\boldsymbol{\delta}) = \boldsymbol{\delta}^{\dagger}\mathbf{M}\boldsymbol{\delta} \text{ such that } \boldsymbol{\delta} \in \Omega, \tag{4.59}$$

where Ω is some constraint set. The Hermitian positive-definite matrix $\mathbf{M}$ is given by

$$\mathbf{M} = \mathbf{R}^{\mathrm{T}} \left(\mathbf{I}_{N_{\mathrm{c}}} - \mathbf{P}_{\mathrm{r}}^{\mathrm{T}} \right) \mathbf{R}^{*}, \tag{4.60}$$

where $\mathbf{R}$ is a column-wise circulant matrix with the first column vector $\mathbf{r}$. The L-dimensional orthogonal projection matrix $\mathbf{P}_{\mathrm{r}} = \mathbf{S}\mathbf{F}_{\mathrm{t}}\mathbf{B}^{-1}\mathbf{F}_{\mathrm{t}}^{\dagger}\mathbf{S}^{\dagger}$ with $\mathbf{B} = \mathbf{F}_{\mathrm{t}}^{\dagger}\mathbf{S}^{\dagger}\mathbf{S}\mathbf{F}_{\mathrm{t}}$.

In [119], the authors propose to use the constraint set given by

$$\Omega = \left\{ \boldsymbol{\delta} \;\middle|\; \frac{1}{2}\left(\boldsymbol{\delta}^{\dagger}\mathbf{e}_1 + \mathbf{e}_1^{\dagger}\boldsymbol{\delta} \right) = 1 \right\}, \tag{4.61}$$

where the $N_{\mathrm{c}} \times 1$ vector $\mathbf{e}_1 = [1, 0, \ldots, 0]^{\mathrm{T}}$. Depending upon the level of phase noise, the constraint set in (4.61) may not necessarily be optimal, i.e., the true phase noise spectral vector may not lie in this set. For example, in the absence of phase noise, we have that $\boldsymbol{\delta} = \mathbf{e}_1$. Thus, for slow-varying phase noise processes, we can expect strong correlation between $\boldsymbol{\delta}$ and $\mathbf{e}_1$ which justifies the use of (4.61). However, it is not applicable for moderately varying or fast-varying phase noise processes.

Irrespective of the rate at which phase noise varies, in Publication VI, possible subspaces in which the vector $\boldsymbol{\delta}$ may lie are explored. The following proposition, originally derived in Publication VI, paves the way for a subspace-based approach.

Proposition 4.3.1. *Denote the null space of* $\mathbf{M}$ *by* $\mathcal{N}(\mathbf{M})$. *Then at infinite SNR,* $\boldsymbol{\delta} \in \mathcal{N}(\mathbf{M})$.

Proof. See Publication VI. □

Proposition 4.3.1 invites for a few remarks. Firstly, the proposition reveals to us about *where to look* for $\boldsymbol{\delta}$ when minimizing $\mathcal{L}(\boldsymbol{\delta})$. Secondly, the proposition is also applicable at high SNRs, i.e., we can expect $\boldsymbol{\delta}$ to be close to the null space of $\mathbf{M}$. Finally, it is useful to know in *how big a space* does $\boldsymbol{\delta}$ lie in, i.e., the dimensionality of $\mathcal{N}(\mathbf{M})$. From (4.60), we can see that, in general, $\mathbf{R}^{*}$ is a full-rank matrix, and since $\left(\mathbf{I}_{N_{\mathrm{c}}} - \mathbf{P}_{\mathrm{r}}^{\mathrm{T}} \right)$ has rank $N_{\mathrm{c}} - L$, we must have the rank of $\mathbf{M}$ also equal to $N_{\mathrm{c}} - L$ and, thus, the dimensionality of $\mathcal{N}(\mathbf{M})$ is equal to L which is number of channel taps.

4.3.1 Subspace-based Minimization Schemes

Before we proceed, we digress to spare a thought about *whether we need to estimate all the components of $\boldsymbol{\delta}$ when minimizing $\mathcal{L}(\boldsymbol{\delta})$*. In general, oscillators are designed such that there is tolerable amount of spectral spreading of the RF-carrier signal. This essentially implies that most of the power of the vector $\boldsymbol{\delta}$ is confined to only a few of its (low frequency) components. Based on this observation, we can reduce the minimization of $\mathcal{L}(\boldsymbol{\delta})$ to only N variables, where $N \ll N_{\mathrm{c}}$.

Denote $\underline{\boldsymbol{\delta}}$ as the $N \times 1$ vector that comprises the N components of $\boldsymbol{\delta}$. We model the relation between the vectors as

$$\boldsymbol{\delta} = \begin{pmatrix} \mathbf{I}_{\frac{N}{2}} & \mathbf{0}_{\frac{N}{2}\times\frac{N}{2}} \\ \mathbf{0}_{(N_{\mathrm{c}}-N)\times\frac{N}{2}} & \mathbf{0}_{(N_{\mathrm{c}}-N)\times\frac{N}{2}} \\ \mathbf{0}_{\frac{N}{2}\times\frac{N}{2}} & \mathbf{I}_{\frac{N}{2}} \end{pmatrix} \underline{\boldsymbol{\delta}} = \mathbf{L}\underline{\boldsymbol{\delta}}. \tag{4.62}$$

From the above model, we see that we only keep the top and bottom part of $\boldsymbol{\delta}$ which corresponds to positive and negative low frequencies centered around zero, while the high-frequency components are set to a value of zero. With this model in place, we can reduce the complexity of minimizing $\mathcal{L}(\boldsymbol{\delta})$ from N_{c} unknowns to only N unknowns. The transformed likelihood function is obtained by substituting (4.62) in $\mathcal{L}(\boldsymbol{\delta})$ to obtain

$$\mathcal{L}(\underline{\boldsymbol{\delta}}) = \underline{\boldsymbol{\delta}}^{\dagger}\tilde{\mathbf{M}}\underline{\boldsymbol{\delta}}, \tag{4.63}$$

where $\tilde{\mathbf{M}} = \mathbf{L}^{\dagger}\mathbf{M}\mathbf{L}$. We are now ready to discuss the subspace-based schemes.

NsPM: Nullspace-based Phase Noise Minimization

We would like use the information that, at infinite SNR, $\boldsymbol{\delta} \in \mathcal{N}(\mathbf{M})$ when minimizing (4.63). Let $\mathbf{N}$ denote the matrix whose columns span $\mathcal{N}(\mathbf{M})$. Since we model $\boldsymbol{\delta}$ by (4.62), requiring $\boldsymbol{\delta} \in \mathcal{N}(\mathbf{M})$ implies requiring $\underline{\boldsymbol{\delta}} \in \mathrm{span}(\mathbf{L}^{\dagger}\mathbf{N})$ (we use the fact that $\mathbf{L}^{\dagger}\mathbf{L} = \mathbf{I}_N$), where $\mathrm{span}(\mathbf{X})$ denotes span of the columns of the matrix $\mathbf{X}$. Thus, the optimization problem can be framed as follows:

$$\begin{aligned} &\text{Minimize } \ \mathcal{L}(\underline{\boldsymbol{\delta}}) = \underline{\boldsymbol{\delta}}^{\dagger}\tilde{\mathbf{M}}\underline{\boldsymbol{\delta}} \\ &\text{such that } \ \underline{\boldsymbol{\delta}}^{\dagger}\underline{\boldsymbol{\delta}} = 1, \ \underline{\boldsymbol{\delta}} \in \mathrm{span}(\mathbf{L}^{\dagger}\mathbf{N}). \end{aligned} \tag{4.64}$$

In (4.64), we have enforced a unit-norm constraint on $\underline{\boldsymbol{\delta}}$. Using Parseval's theorem, it can be easily shown that the norm of $\boldsymbol{\delta}$ is one, and since we assume most of the power is in $\underline{\boldsymbol{\delta}}$ then the unit-norm constraint in (4.64)

is reasonable. Writing $\underline{\boldsymbol{\delta}} = \mathbf{L}^\dagger \mathbf{N} \boldsymbol{\alpha}$, the above problem expressed in terms of $\boldsymbol{\alpha}$ is given by

$$\begin{aligned} &\text{Minimize } \; C(\boldsymbol{\alpha}) = \boldsymbol{\alpha}^\dagger \left(\mathbf{N}^\dagger \mathbf{L} \tilde{\mathbf{M}} \mathbf{L}^\dagger \mathbf{N} \right) \boldsymbol{\alpha} \\ &\text{such that } \; \boldsymbol{\alpha}^\dagger \left(\mathbf{N}^\dagger \mathbf{L} \mathbf{L}^\dagger \mathbf{N} \right) \boldsymbol{\alpha} = 1. \end{aligned} \tag{4.65}$$

The optimization problem (4.65) can be solved as follows: Let $\mathbf{N}^\dagger \mathbf{L} \mathbf{L}^\dagger \mathbf{N} = \mathbf{W} \mathbf{W}^\dagger$ be the Cholesky decomposition. Then writing $\boldsymbol{\gamma} = \mathbf{W}^\dagger \boldsymbol{\alpha}$, the minimization problem simplifies to

$$\text{Minimize } \; C(\boldsymbol{\gamma}) = \boldsymbol{\gamma}^\dagger \mathbf{Q} \boldsymbol{\gamma} \text{ such that } \boldsymbol{\gamma}^\dagger \boldsymbol{\gamma} = 1, \tag{4.66}$$

where $\mathbf{Q} = (\mathbf{W}^{-1}) \mathbf{N}^\dagger \mathbf{L} \tilde{\mathbf{M}} \mathbf{L}^\dagger \mathbf{N} (\mathbf{W}^\dagger)^{-1}$. The minimum value for the above problem is equal to the smallest eigenvalue of $\mathbf{Q}$, and if the eigenvalues are distinct then the minimizer corresponds to the eigenvector associated with the smallest eigenvalue.

CvPM: Covariance-based Phase Noise Minimization

Subspace information about $\boldsymbol{\delta}$ can also be obtained from its covariance matrix. It is shown in [129, Appendix C] that a random vector $\mathbf{x}$ will always be drawn from the space spanned by the eigenvectors of its covariance matrix. Let $\mathbf{C}_{\boldsymbol{\delta}}$ denote the covariance matrix of $\boldsymbol{\delta}$. Closed-form expressions of $\mathbf{C}_{\boldsymbol{\delta}}$ for Wiener and PLL-type phase noise processes can be found in [110]. Using the model in (4.62), the covariance matrix of $\underline{\boldsymbol{\delta}}$ is $\mathbf{C}_{\underline{\boldsymbol{\delta}}} = \mathbf{L}^\dagger \mathbf{C}_{\boldsymbol{\delta}} \mathbf{L}$. With these definitions in place, we can frame the following covariance based optimization problem:

$$\text{Minimize } \mathcal{L}(\underline{\boldsymbol{\delta}}) = \underline{\boldsymbol{\delta}}^\dagger \tilde{\mathbf{M}} \underline{\boldsymbol{\delta}}, \text{ s.t } \; \underline{\boldsymbol{\delta}}^\dagger \underline{\boldsymbol{\delta}} = 1, \; \underline{\boldsymbol{\delta}} \in \text{span}(\mathbf{U}), \tag{4.67}$$

where $N \times N$ unitary matrix $\mathbf{U}$ contains the eigenvectors of $\mathbf{C}_{\underline{\boldsymbol{\delta}}}$. Making a variable change by writing $\underline{\boldsymbol{\delta}} = \mathbf{U} \alpha$ and noting that $\mathbf{U}^\dagger \mathbf{U} = \mathbf{I}_N$, we have

$$\text{Minimize } \mathcal{L}(\boldsymbol{\alpha}) = \boldsymbol{\alpha}^\dagger (\mathbf{U}^\dagger \tilde{\mathbf{M}} \mathbf{U}) \boldsymbol{\alpha}, \;\; \text{s.t } \; \boldsymbol{\alpha}^\dagger \boldsymbol{\alpha} = 1. \tag{4.68}$$

The minimizer is equal to the eigenvector associated with the smallest eigenvalue of $(\mathbf{U}^\dagger \tilde{\mathbf{M}} \mathbf{U})$.

CoPM: Correlation-based Phase Noise Minimization of [119]

We now compare the minimization scheme proposed in [119] with the subspace-based approaches. Specifically, the optimization problem considered in [119] is given by

$$\text{Minimize } \mathcal{L}(\underline{\boldsymbol{\delta}}) = \underline{\boldsymbol{\delta}}^\dagger \tilde{\mathbf{M}} \underline{\boldsymbol{\delta}} \;\; \text{s.t } \; \frac{1}{2} \left(\underline{\boldsymbol{\delta}}^\dagger \mathbf{L}^\dagger \mathbf{e}_1 + \mathbf{e}_1^\dagger \mathbf{L} \underline{\boldsymbol{\delta}} \right) = 1, \tag{4.69}$$

where $\mathbf{e}_1 = [1,\ 0, \ldots, 0]^{\mathrm{T}}$ is a $N_{\mathrm{c}} \times 1$ column vector. As previously mentioned, the constraint in (4.69) requires that the minimizer have maximum correlation with the vector $\mathbf{e}_1$. In the absence of phase noise, the actual phase noise spectral vector $\boldsymbol{\delta} = \mathbf{L}\mathbf{e}_1$. For very small phase noise levels, we can expect $\boldsymbol{\delta}$ to be very close to $\mathbf{L}\mathbf{e}_1$ and, thus, the constraint in (4.69) is applicable in this case. However, for fast-varying phase noise processes, the correlation between $\boldsymbol{\delta}$ and $\mathbf{L}\mathbf{e}$ gets weaker and, thereby, using the constraint in (4.69) will yield poor phase noise estimates. The minimizer to (4.69) can be easily derived and is given by

$$\underline{\boldsymbol{\delta}} = \frac{\tilde{\mathbf{M}}^{-1}\mathbf{L}\mathbf{e}_1}{\mathbf{e}_1^{\dagger}\left(\mathbf{L}^{\dagger}\tilde{\mathbf{M}}^{-1}\mathbf{L}\right)\mathbf{e}_1}. \tag{4.70}$$

4.3.2 Numerical Results

Let us now present some numerical results on the subspace-based phase noise estimation schemes. A detailed discussion on these results can be found Publication VI. Specifically, we compare the performance of the NsPM, CvPM and CoPM schemes, wherein the performance metric used is the mean-square-error (MSE). The MSE is obtained by calculating the error between the estimate obtained from solving the minimization schemes in the previous section and the true vector phase noise spectral vector. The results are then averaged over many realizations of the preamble OFDM symbol. The channel MSE is evaluated by calculating the error between $\hat{\mathbf{h}}_{\mathrm{ML}}$ of (4.58) and the true channel vector $\mathbf{h}$. As can be seen from (4.58), the channel estimate requires knowledge of $\boldsymbol{\delta}$ through the $\mathbf{V}$ matrix. Once an estimate of $\boldsymbol{\delta}$ is obtained, the columns of $\mathbf{V}$ are formed by circularly shifting $\boldsymbol{\delta}$. Thus, we see that the channel MSE is in direct correspondence with the phase noise MSE.

The system parameters used in the simulations are as follows: The number of subcarriers $N_{\mathrm{c}} = 512$, subcarrier spacing $f_{\mathrm{sub}} = 15$ kHz and bandwidth is equal to 7.7 MHz. For phase noise estimation, we estimate a total of $N = 7$ components of $\boldsymbol{\delta}$ while the rest are set to a value of zero. The symbol constellation is 16-QAM. The channel is Rayleigh fading with exponential power delay profile and number of taps (L) set to four, i.e., $L = 4$. The quantity $\rho = \frac{f_{3\mathrm{dB}}}{f_{\mathrm{sub}}}$ denotes the normalized phase noise 3-dB bandwidth.

Figure 4.3 shows phase noise and channel MSE plots for the NsPM, CvPM and CoPM schemes. The first observation from these plots is that

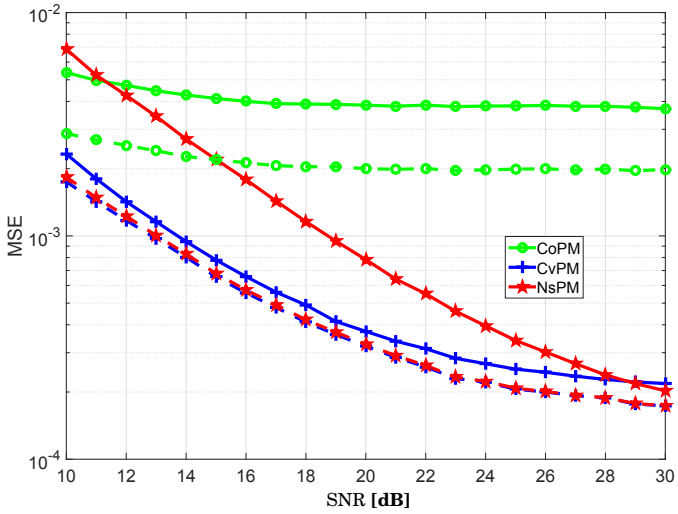


Figure 4.3. Phase noise and channel MSE as a function of the signal-to-noise ratio. The phase noise MSE curves are shown by the solid lines while the dashed lines are the channel MSE curves. The value of $\rho = 0.02$.

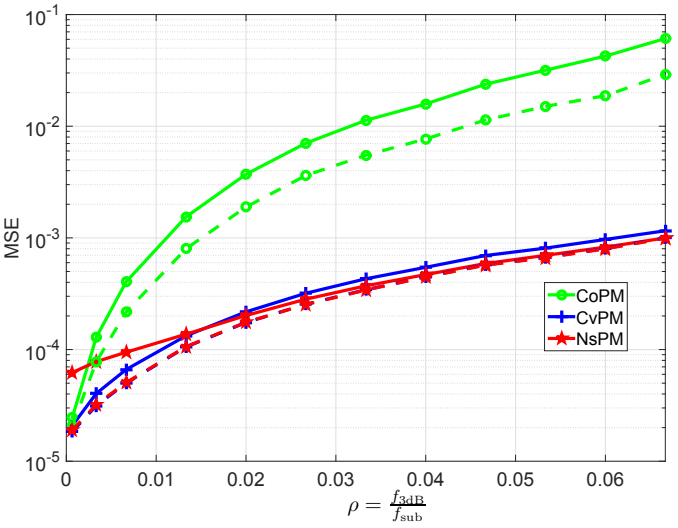


Figure 4.4. Phase noise and channel MSE as a function of the ratio $\rho = \frac{f_{3\text{dB}}}{f_{\text{sub}}}$. The phase noise MSE curves are shown by the solid lines while the dashed lines are the channel MSE curves. The SNR is set to 30 dB.

the subspace-based approaches yield superior MSE performance compared to the CoPM scheme of [119]. Among the subspace-based methods, CvPM performs best. This is expected as second order statistical information is used in the form of the eigenspace of the covariance matrix of δ which is

applicable at any SNR. The NsPM scheme works on the premise that $\boldsymbol{\delta}$ always lie in the null space of M which is true only at infinite SNR. Thus, at low and medium SNRs, we can expect inferior performance when compared to CvPM as witnessed in the figure. At higher SNRs, however, the MSE values of NsPM are similar to that of CvPM.

In Fig. 4.4, we investigate the behavior of the estimations schemes by increasing the normalized 3-dB bandwidth, i.e., the parameter ρ. A small value of ρ effectively implies a small level of phase noise experienced by the OFDM receiver and vice versa. As seen from the figure, the subspace schemes demonstrate superior MSE performance compared to CoPM. This is expected because, the subspace-based schemes utilize subspace information which does not depend upon the set phase noise level. On the other hand CoPM requires maximum correlation with the vector $\mathbf{e}_1$ which is true only for slow-varying phase noise processes.

4.4 Geometry-based Phase Noise Estimation of Publication V

The work in Publication VI demonstrated the use of subspace information when minimizing the cost function $\mathcal{L}(\boldsymbol{\delta})$. In Publication V, information on the *geometry of* $\boldsymbol{\delta}$ is utilized when minimizing $\mathcal{L}(\boldsymbol{\delta})$. The rest of this section is a summary of the work originally described in Publication V. In this thesis, we refer to $\boldsymbol{\delta}$ as the 'phase noise spectral vector'.

4.4.1 Geometry of the Phase Noise Spectral Vector

The vector $\boldsymbol{\delta}$ is the DFT of $\boldsymbol{\phi} = \frac{1}{N_\mathrm{c}}\left[e^{\jmath\theta[0]},\ e^{\jmath\theta[1]}, \ldots, e^{\jmath\theta[N_\mathrm{c}-1]}\right]^\mathrm{T}$ whose elements always have *constant magnitude*. Thus, one can suspect that this constant magnitude property in the time domain must appear in some equivalent form in the frequency domain also. Consider the equation:

$$\left(\frac{e^{\jmath\theta[n]}}{N_\mathrm{c}}\right)\left(\frac{e^{-\jmath\theta[n]}}{N_\mathrm{c}}\right) = \frac{1}{N_\mathrm{c}^2}. \tag{4.71}$$

Taking the N_c-point DFT on both sides of (4.71), and using the fact that N_c-point $\mathrm{DFT}\left[\frac{e^{\jmath\theta[n]}}{N_\mathrm{c}}\right] = \delta_k$ implies $\mathrm{DFT}\left[\frac{e^{-\jmath\theta[n]}}{N_\mathrm{c}}\right] = \delta^*_{-k}$ and $\mathrm{DFT}\left[x[n]y[n]\right] = \frac{1}{N_\mathrm{c}} X_k \circledast Y_k$, where $\circledast$ denotes circular convolution, we arrive at

$$\sum_{k=0}^{N_\mathrm{c}-1} \delta_k \delta^*_{k+l} = \Lambda_l,\ l = 0, 1, 2, \ldots, N_\mathrm{c} - 1, \tag{4.72}$$

where Λ_l is the Kronecker delta function, i.e., $\Lambda_0 = 1$ and $\Lambda_l = 0$ for $l > 0$. We can express (4.72) more compactly as

$$\boldsymbol{\delta}^{\dagger}\mathbf{P}_l\boldsymbol{\delta} = \Lambda_l,\ l = 0, 1, \ldots, N_{\mathrm{c}} - 1, \tag{4.73}$$

where $\mathbf{P}_l = (\mathbf{P}_1)^l$ is a permutation matrix defined by the $N_{\mathrm{c}} \times N_{\mathrm{c}}$ matrix $\mathbf{P}_1$. The first column of $\mathbf{P}_1$ is given by the $N_{\mathrm{c}} \times 1$ vector $[0, 1, 0, \ldots, 0]^{\mathrm{T}}$ and the jth column is obtained by circularly shifting the vector $j-1$ times to the bottom. For $l = 0$, we get the unit-norm property of $\boldsymbol{\delta}$, where $\mathbf{P}_0 = \mathbf{I}_{N_{\mathrm{c}}}$ is the identity matrix.

4.4.2 Geometry-preserving Dimensionality Reduction

In Section 4.3.1, in order to reduce the complexity when minimizing $\mathcal{L}(\boldsymbol{\delta})$, a linear model is used to describe $\boldsymbol{\delta}$ using a transformation matrix $\mathbf{L}$ and a smaller N-dimensional vector $\underline{\boldsymbol{\delta}}$. These are related as

$$\boldsymbol{\delta} = \begin{pmatrix} \mathbf{I}_{\frac{N}{2}} & \mathbf{0}_{\frac{N}{2}\times\frac{N}{2}} \\ \mathbf{0}_{(N_{\mathrm{c}}-N)\times\frac{N}{2}} & \mathbf{0}_{(N_{\mathrm{c}}-N)\times\frac{N}{2}} \\ \mathbf{0}_{\frac{N}{2}\times\frac{N}{2}} & \mathbf{I}_{\frac{N}{2}} \end{pmatrix} \underline{\boldsymbol{\delta}} = \mathbf{L}\underline{\boldsymbol{\delta}}. \tag{4.74}$$

At the same time, we would also like to utilize the geometry of $\boldsymbol{\delta}$ when minimizing $\mathcal{L}(\boldsymbol{\delta})$. Using (4.74) in (4.73), the minimization problem is framed as

$$\begin{aligned} &\text{Minimize}\ \ \mathcal{L}(\underline{\boldsymbol{\delta}}) = \underline{\boldsymbol{\delta}}^{\dagger}\left(\mathbf{L}^{\dagger}\mathbf{M}\mathbf{L}\right)\underline{\boldsymbol{\delta}} \\ &\text{s.t}\ \ \underline{\boldsymbol{\delta}}^{\dagger}\left(\mathbf{L}^{\dagger}\mathbf{P}_l\mathbf{L}\right)\underline{\boldsymbol{\delta}} = \Lambda_l,\ l = 0, 1, \ldots, N_{\mathrm{c}} - 1. \end{aligned} \tag{4.75}$$

One advantage of solving the above optimization problem is that the number of unknowns is reduced from N_{c} to N. However, the number of constraints, which is still equal to N_{c}, can cause the complexity to be very high. For example, OFDM systems with $N_{\mathrm{c}} = 8192$ subcarriers, solving (4.75) can be computationally high even if N is much less than N_{c}.

In conclusion, using (4.74) in tandem with (4.73) may not render a feasible and practical optimization problem from the point of view of computational complexity. To achieve a less complex minimization scheme that also utilizes the phase noise geometry requires *modifying the linear model to incorporate this knowledge of phase noise geometry*. This is the topic of the next subsection.

Phase Noise Geometry Preserving Transformation (PPT)

We would like to model $\boldsymbol{\delta}$ using a linear model, i.e.,

$$\boldsymbol{\delta} = \mathbf{T}\underline{\boldsymbol{\delta}}, \tag{4.76}$$

where the $N_\mathrm{c} \times N$ matrix $\mathbf{T}$ represents the transformation from the smaller N-dimensional space to the bigger N_c-dimensional space. At the same time, we would like $\boldsymbol{\delta}$ in (4.76) to satisfy (4.73). *The ideology of this modeling is that all information about $\boldsymbol{\delta}$ can be equivalently acquired from its smaller subset $\underline{\boldsymbol{\delta}}$. In line with this, we can think of $\boldsymbol{\delta}$ acquiring its geometry from $\underline{\boldsymbol{\delta}}$ which is a N-dimensional phase noise spectral vector that satisfies the phase noise geometry in the smaller space.* The N-dimensional equivalent of (4.73) is given by

$$\underline{\boldsymbol{\delta}}^\dagger \tilde{\mathbf{P}}_l \underline{\boldsymbol{\delta}} = \tilde{\Lambda}_l, \; l = 0, 1, \ldots, N-1, \tag{4.77}$$

where $\tilde{\mathbf{P}}_l$ and $\tilde{\Lambda}_l$ are the N-dimensional equivalents of $\mathbf{P}_l$ and Λ_l respectively. Thus, when moving from N-dimensional to the N_c-dimensional space, the matrix $\mathbf{T}$ is chosen such that the phase noise geometry is also preserved in the N_c-dimensional space. Given that $\underline{\boldsymbol{\delta}}$ satisfies (4.77), not all transformation matrices $\mathbf{T}$ allow $\boldsymbol{\delta}$ in (4.76) to satisfy (4.73). In Publication V, this problem is investigated, and it is shown that there exists transformation matrices that preserve this geometry and also closed-form expressions for the columns of such matrices are given. Such transformation matrices are referred to as *phase noise geometry preserving transformations (PPTs)*. The generic form of a PPT is given by

$$\mathbf{T} = \mathbf{F}\tilde{\mathbf{T}}\tilde{\mathbf{F}}^\dagger, \tag{4.78}$$

where the respective matrices $\mathbf{F}$ and $\tilde{\mathbf{F}}$ are the $N_\mathrm{c} \times N_\mathrm{c}$ and $N \times N$ DFT matrices. The columns $\tilde{\mathbf{t}}_i$ of the $N_\mathrm{c} \times N$ matrix $\tilde{\mathbf{T}}$ must satisfy, for all $l = 1, 2, \ldots, N_\mathrm{c} - 1$,

$$\tilde{\mathbf{T}}^\dagger \tilde{\mathbf{T}} = \mathbf{I}_N, \; \tilde{\mathbf{t}}_i^\dagger \mathbf{D}_l \tilde{\mathbf{t}}_j = 0 \text{ for } i \neq j, \; \sum_{i=0}^{N-1} \tilde{\mathbf{t}}_i^\dagger \mathbf{D}_l \tilde{\mathbf{t}}_i = 0, \tag{4.79}$$

where the $N_\mathrm{c} \times N_\mathrm{c}$ diagonal matrix $\mathbf{D}_l = \mathbf{F}^\dagger \mathbf{P}_l \mathbf{F}$ and comprises of the eigenvalues of the permutation matrix $\mathbf{P}_l$.

The Piecewise Constant PPT (PC-PPT)

We now present one simple example of a PPT. The general form of PPTs in (4.78) renders a nice interpretation. The transformation $\boldsymbol{\delta} = \mathbf{F}\tilde{\mathbf{T}}\tilde{\mathbf{F}}^\dagger \underline{\boldsymbol{\delta}}$ can be viewed as interpolation of the smaller N-dimensional signal $\tilde{\mathbf{F}}^\dagger \underline{\boldsymbol{\delta}}$ to a higher N_c-dimensional vector. The interpolation is performed by the $\tilde{\mathbf{T}}$ matrix. The result of the interpolation is then transformed to the Fourier domain by the $\mathbf{F}$ matrix. Phase noise processes in general are lowpass processes and, thus, such an interpretation is valid. One of the simplest

interpolators is to simply repeat the elements of the time-domain vector, i.e.,

$$\mathbf{T}_{\mathrm{PC}} = \sqrt{\frac{N_c}{N}}\mathbf{F}\begin{pmatrix} \mathbf{1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{1} \end{pmatrix}\tilde{\mathbf{F}}^{\dagger}, \tag{4.80}$$

where $\mathbf{1}$ is a column vector of length $\frac{N_c}{N}$ with all elements equal to one, and $\mathbf{0}$ denotes the vector with all elements equal to zero. It can be easily verified that the above matrix indeed satisfies the conditions in (4.79) and, hence, is a PPT.

4.4.3 Geometry-based Minimization Schemes

In this section, we review some phase noise optimization problems originally proposed in Publication V. We begin with the optimal approach where we explicitly impose the geometry of $\boldsymbol{\delta}$ when minimizing $\mathcal{L}(\boldsymbol{\delta})$. Later on, sub-optimal schemes, however of lower computational complexity, are reviewed.

Phase Noise Constraints (PNC)

We minimize $\mathcal{L}(\boldsymbol{\delta})$ using the linear model in (4.76) and impose (4.77) as constraints. Specifically, the optimization problem is given by

$$\begin{aligned} (\mathcal{S}):\ &\text{Minimize } \mathcal{L}(\underline{\boldsymbol{\delta}}) = \underline{\boldsymbol{\delta}}^{\dagger}\tilde{\mathbf{M}}\underline{\boldsymbol{\delta}} \\ &\text{s.t } \ \underline{\boldsymbol{\delta}}^{\dagger}\underline{\boldsymbol{\delta}} = 1,\ \underline{\boldsymbol{\delta}}^{\dagger}\tilde{\mathbf{P}}_l\underline{\boldsymbol{\delta}} = 0,\ l = 1, 2, \dots, K-1, \end{aligned} \tag{4.81}$$

where $\tilde{\mathbf{M}} = \mathbf{T}^{\dagger}\mathbf{M}\mathbf{T}$. In the above problem, the number of constraints are parametrized by K. This allows flexibility in choosing the number of constraints. Note that it is sufficient that $K \leq \frac{N+1}{2}$, since $\left(\underline{\boldsymbol{\delta}}^{\dagger}\tilde{\mathbf{P}}_l\underline{\boldsymbol{\delta}}\right)^{\dagger} = 0^{\dagger}$ implies $\underline{\boldsymbol{\delta}}^{\dagger}\tilde{\mathbf{P}}_{N-l}\underline{\boldsymbol{\delta}} = 0$ after using $\tilde{\mathbf{P}}_l^{\dagger} = \tilde{\mathbf{P}}_{N-l}$. In general, the quadratic form $\underline{\boldsymbol{\delta}}^{\dagger}\tilde{\mathbf{P}}_l\underline{\boldsymbol{\delta}}$ yields complex values as the eigenvalues of $\tilde{\mathbf{P}}_l, l > 0$ are complex-valued. Thus, the constraint in (4.81) can be expressed in terms of its real an imaginary parts as

$$\begin{aligned} (\mathcal{S}):\ &\text{Minimize } \mathcal{L}(\underline{\boldsymbol{\delta}}) = \underline{\boldsymbol{\delta}}^{\dagger}\tilde{\mathbf{M}}\underline{\boldsymbol{\delta}} \\ &\text{s.t } \ \underline{\boldsymbol{\delta}}^{\dagger}\underline{\boldsymbol{\delta}} = 1,\ \underline{\boldsymbol{\delta}}^{\dagger}\tilde{\mathbf{P}}_l^{R}\underline{\boldsymbol{\delta}} = 0,\ \underline{\boldsymbol{\delta}}^{\dagger}\tilde{\mathbf{P}}_l^{I}\underline{\boldsymbol{\delta}} = 0,\ l = 1, 2, \dots, K-1, \end{aligned} \tag{4.82}$$

where $\tilde{\mathbf{P}}_l^{R} = \frac{\tilde{\mathbf{P}}_l + \tilde{\mathbf{P}}_l^{\dagger}}{2}$ and $\tilde{\mathbf{P}}_l^{I} = \frac{\jmath(\tilde{\mathbf{P}}_l^{\dagger} - \tilde{\mathbf{P}}_l)}{2}$ represent the real and imaginary parts of $\tilde{\mathbf{P}}_l$.

Problems such as that of $(\mathcal{S})$ come under the subject of *optimization theory* [130]. Important questions such as existence of global minimum

and solvability of the problem in polynomial time are addressed using optimization theory. In Publication V, it is shown that the global minimum can be obtained by equivalently solving the *convex* dual problem to $(\mathcal{S})$ [130]. Convex problems are guaranteed to yield global solutions in polynomial time.

The rest of this subsection is a digression to optimization theory, where we try to highlight how the optimal solution of (4.82) can be attained. In going through this part, the curious reader will be introduced to the wonderful topic called *S-procedure*. Of course, the reader may skip this part and proceed directly to the subsection on *'Suboptimal Schemes'* without any loss in continuity.

Optimality of $(\mathcal{S})$***:*** In the language of optimization theory, $(\mathcal{S})$ is referred to as the *primal problem* which is a *non-convex program* [130]: The cost function is a convex function because $\tilde{\mathbf{M}}$ is a positive-definite Hermitian matrix. However, the constraint functions are non-convex because $\tilde{\mathbf{P}}_l^R$ and $\tilde{\mathbf{P}}_l^I$ are indefinite matrices. In general, if $(\mathcal{S})$ is a *convex program* then every local minimum is also a global minimum which eases the task in finding the global solution. All the more, convex programs can be solved in polynomial time using interior-point algorithms [131]. These nice properties, however, are not necessarily satisfied by non-convex problems. A suboptimal solution to $(\mathcal{S})$ can be obtained by solving the so-called *dual problem*. The corresponding dual problem to $(\mathcal{S})$ can be easily derived and is given by [130]

$$\begin{aligned}(\mathcal{DS}):\ &\text{Maximize}\ \ \lambda \\ &\text{s.t}\ \ \tilde{\mathbf{M}} - \lambda \mathbf{I}_N + \sum_{l=1}^{K-1} \alpha_l \tilde{\mathbf{P}}_l^R + \beta_l \tilde{\mathbf{P}}_l^I \succeq 0, \end{aligned} \tag{4.83}$$

where optimization is done over the variables λ, α_l and β_l. The dual problem $(\mathcal{DS})$ is always a convex program and yields a solution that is always less than or equal to the optimal value of the primal problem. The difference in value of the solution of $(\mathcal{S})$ and $(\mathcal{DS})$ is called the *duality gap*. When the duality gap is zero, also known as strong duality, it implies solving the convex dual problem is equivalent to solving the original primal problem. Hence, with zero-duality gap, even when $(\mathcal{S})$ is non-convex, the optimal solution is still attained by solving the convex dual problem.

In Publication V, it is shown that the strong duality holds between $(\mathcal{S})$ and $(\mathcal{DS})$. We now shed some light on how this might be possible. First,

we define the following:

$$s_0(\underline{\boldsymbol{\delta}}) = \begin{pmatrix} \underline{\boldsymbol{\delta}} \\ 1 \end{pmatrix}^{\dagger} \begin{pmatrix} \tilde{\mathbf{M}} & \\ & -\gamma \end{pmatrix} \begin{pmatrix} \underline{\boldsymbol{\delta}} \\ 1 \end{pmatrix}, s_1(\underline{\boldsymbol{\delta}}) = \begin{pmatrix} \underline{\boldsymbol{\delta}} \\ 1 \end{pmatrix}^{\dagger} \begin{pmatrix} \mathbf{I}_N & \\ & -1 \end{pmatrix} \begin{pmatrix} \underline{\boldsymbol{\delta}} \\ 1 \end{pmatrix}; \quad (4.84)$$

$$u_l(\underline{\boldsymbol{\delta}}) = \begin{pmatrix} \underline{\boldsymbol{\delta}} \\ 1 \end{pmatrix}^{\dagger} \begin{pmatrix} \tilde{\mathbf{P}}_l^R & \\ & 0 \end{pmatrix} \begin{pmatrix} \underline{\boldsymbol{\delta}} \\ 1 \end{pmatrix}, v_l(\underline{\boldsymbol{\delta}}) = \begin{pmatrix} \underline{\boldsymbol{\delta}} \\ 1 \end{pmatrix}^{\dagger} \begin{pmatrix} \tilde{\mathbf{P}}_l^I & \\ & 0 \end{pmatrix} \begin{pmatrix} \underline{\boldsymbol{\delta}} \\ 1 \end{pmatrix}, \quad (4.85)$$

where $\underline{\boldsymbol{\delta}} \in \mathcal{C}^N$. Define the sets

$$\Pi = \left\{ \Big(s_0(\underline{\boldsymbol{\delta}}),\ s_1(\underline{\boldsymbol{\delta}}),\ u_1(\underline{\boldsymbol{\delta}}), \ldots, u_{K-1}(\underline{\boldsymbol{\delta}}),\ v_1(\underline{\boldsymbol{\delta}}), \ldots, v_{K-1}(\underline{\boldsymbol{\delta}}) \Big)^{\mathrm{T}} : \underline{\boldsymbol{\delta}} \in \mathcal{C}^N \right\}, \quad (4.86)$$

$$\mathcal{N} = \left\{ \Big(g, \mathbf{0}_{2(K-1)+1}^{\mathrm{T}} \Big)^{\mathrm{T}} \text{ s.t } g < 0 \right\}, \quad (4.87)$$

where $\mathbf{0}_{2(K-1)+1}$ is the $2(K-1)+1 \times 1$ vector of zeros. The optimization problem $(\mathcal{S})$ can be reformulated as

$$(\mathcal{S}): \quad \text{Maximize } \gamma \quad \text{s.t } \underline{\boldsymbol{\delta}}^{\dagger} \tilde{\mathbf{M}} \underline{\boldsymbol{\delta}} \geq \gamma \text{ for all } \underline{\boldsymbol{\delta}} \in \Upsilon, \quad (4.88)$$

where Υ is the feasible set of (4.82), i.e., $\Upsilon = \{\underline{\boldsymbol{\delta}} \mid \underline{\boldsymbol{\delta}}^{\dagger} \underline{\boldsymbol{\delta}} = 1,\ \underline{\boldsymbol{\delta}}^{\dagger} \tilde{\mathbf{P}}_l^R \underline{\boldsymbol{\delta}} = 0,\ \underline{\boldsymbol{\delta}}^{\dagger} \tilde{\mathbf{P}}_l^I \underline{\boldsymbol{\delta}} = 0, l = 1, 2, \ldots, K-1\}$. The constraint in (4.88) is the same as saying $s_0(\underline{\boldsymbol{\delta}}) \geq 0$ whenever $s_1(\underline{\boldsymbol{\delta}}) = 0$, $u_i(\underline{\boldsymbol{\delta}}) = 0$ and $v_i(\underline{\boldsymbol{\delta}}) = 0$ for all $i = 1, 2, \ldots, K-1$. Equivalently, when expressed in terms of the sets Π and $\mathcal{N}$, it implies $\Pi \cap \mathcal{N} = \emptyset$, where $\emptyset$ denotes the empty set. Thus, the optimization problem can be rewritten as

$$(\mathcal{S}): \quad \text{Maximize } \gamma \text{ such that } \Pi \cap \mathcal{N} = \emptyset. \quad (4.89)$$

By introducing an auxiliary variable, the dual problem in (4.83) can be rewritten as

$$(\mathcal{DS}): \text{ Maximize } \gamma$$
$$\begin{pmatrix} \tilde{\mathbf{M}} - \lambda \mathbf{I}_N + \sum_{l=1}^{K-1} \alpha_l \tilde{\mathbf{P}}_l^R + \beta_l \tilde{\mathbf{P}}_l^I & \\ & \lambda - \gamma \end{pmatrix} \succeq 0, \quad (4.90)$$

where γ is the auxiliary variable. Comparing (4.89) and (4.90), we see that $(\mathcal{S})$ is exactly the same as $(\mathcal{DS})$ if the condition $\Pi \cap \mathcal{N} = \emptyset$ in (4.89) is equivalent to the matrix inequality in (4.90), and hence, solving either the dual or primal problem will always yield the same value. In the following paragraphs, we show that this equivalence is indeed true. The key to this revelation is the *S-procedure*.

S-procedure: The S-procedure is a method wherein a set of quadratic constraints are replaced by a linear matrix inequality (LMI) [132]. For example, in (4.88), we have a set of quadratic constraints while the constraint in (4.90) is a LMI. To precisely describe it, we first define the following:

$$\sigma_0(\mathbf{x}) = \mathbf{x}^\dagger \begin{pmatrix} \tilde{\mathbf{M}} & \\ & -\gamma \end{pmatrix} \mathbf{x}, \sigma_1(\mathbf{x}) = \mathbf{x}^\dagger \begin{pmatrix} \mathbf{I}_N & \\ & -1 \end{pmatrix} \mathbf{x}; \tag{4.91}$$

$$\psi_l(\mathbf{x}) = \mathbf{x}^\dagger \begin{pmatrix} \tilde{\mathbf{P}}_l^R & \\ & 0 \end{pmatrix} \mathbf{x}, \upsilon_l(\mathbf{x}) = \mathbf{x}^\dagger \begin{pmatrix} \tilde{\mathbf{P}}_l^I & \\ & 0 \end{pmatrix} \mathbf{x}, \tag{4.92}$$

where $\mathbf{x} \in \mathcal{C}^{N+1}$, $l = 1, 2, \ldots, K-1$ and the above matrices are block-diagonal. Define the set

$$\mathcal{Y} = \left\{ \Big(\sigma_0(\mathbf{x}),\ \sigma_1(\mathbf{x}),\ \psi_1(\mathbf{x}), \ldots, \psi_{K-1}(\mathbf{x})\ \upsilon_1(\mathbf{x}), \ldots, \upsilon_{K-1}(\mathbf{x}) \Big)^{\mathrm{T}} : \mathbf{x} \in \mathcal{C}^{N+1} \right\}. \tag{4.93}$$

Consider the following two statements:

- S1: $\sigma_0(\mathbf{x}) \geq 0$ whenever $\sigma_1(\mathbf{x}) = 0,\ \psi_l(\mathbf{x}) = 0,\ \upsilon_l(\mathbf{x}) = 0$ for all $l = 1, 2, \ldots, K-1$. Another way of stating it is $\mathcal{Y} \cap \mathcal{N} = \emptyset$, where $\mathcal{N}$ is defined in (4.87).

- S2: There exists constants λ, α_l and β_l such that

$$\begin{pmatrix} \tilde{\mathbf{M}} - \lambda \mathbf{I}_N + \sum_{l=1}^{K-1} \alpha_l \tilde{\mathbf{P}}_l^R + \beta_l \tilde{\mathbf{P}}_l^I & \\ & \lambda - \gamma \end{pmatrix} \succeq 0. \tag{4.94}$$

We now ask the following question: Are the statements S1 and S2 equivalent ? It can inferred that S2 implies S1. This is seen as follows: Equation (4.94) implies, for all $\mathbf{x} \in \mathcal{C}^{N+1}$, we have

$$\mathbf{x}^\dagger \begin{pmatrix} \tilde{\mathbf{M}} & \\ & -\gamma \end{pmatrix} \mathbf{x} - \lambda \mathbf{x}^\dagger \begin{pmatrix} \mathbf{I}_N & \\ & -1 \end{pmatrix} \mathbf{x} + \sum_{l=1}^{K-1} \alpha_l \mathbf{x}^\dagger \begin{pmatrix} \tilde{\mathbf{P}}_l^R & \\ & 0 \end{pmatrix} \mathbf{x} + \sum_{l=1}^{K-1} \beta_l \mathbf{x}^\dagger \begin{pmatrix} \tilde{\mathbf{P}}_l^I & \\ & 0 \end{pmatrix} \mathbf{x} \geq 0 \tag{4.95}$$

$$\sigma_0(\mathbf{x}) - \lambda \sigma_1(\mathbf{x}) + \sum_{l=1}^{K-1} \alpha_l \psi_l(\mathbf{x}) + \beta_l \upsilon_l(\mathbf{x}) \geq 0 \tag{4.96}$$

$$\mathbf{a}^{\mathrm{T}} \mathbf{y} \geq 0, \mathbf{y} \in \mathcal{Y}, \tag{4.97}$$

where the $2K \times 1$ column vector $\mathbf{a} = [1,\ -\lambda,\ \alpha_1, \ldots, \alpha_{K-1},\ \beta_1, \ldots, \beta_{K-1}]^{\mathrm{T}}$. Equation (4.97) implies that all points of the set $\mathcal{Y}$ always lie on one side of the hyperplane determined by the vector $\mathbf{a}$. Now, for points in the set

$\mathcal{N}$ defined in (4.87), we have

$$\mathbf{a}^{\mathrm{T}}\mathbf{y} = g < 0,\ \mathbf{y} \in \mathcal{N}, \tag{4.98}$$

where the inequality follows since $g < 0$ as defined in (4.87). The above equation in conjunction with (4.97) imply that $\mathcal{Y} \cap \mathcal{N} = \emptyset$.

As we have seen in the previous paragraph, the statement S2 invokes the existence of the statement S1. Unfortunately, S1 does not necessarily imply S2, and only in certain cases and depending upon the problem at hand, the implication may follow. We say that the S-procedure is *lossless* when S1 implies S2, where S2 implies S1 is always understood to be true. In [133], Yakobovich provided conditions on when the S-procedure is lossless for the case with quadratic *inequality* constraints. In our case, in the statement S1, we have quadratic equality constraints. An excellent summary of the S-procedure for inequality constraints is given in [134].

In Publication V, for quadratic forms with equality constraints, conditions for the S-procedure to be lossless are derived. We refer the reader to Publication V for the details. The following theorem summarizes the result before which we need the *regularity condition*.

Define the vector

$$\mathbf{w}(\mathbf{x}) = \Big(\sigma_1(\mathbf{x}), \psi_1(\mathbf{x}), \ldots, \psi_{K-1}(\mathbf{x}), \upsilon_1(\mathbf{x}), \ldots, \upsilon_{K-1}(\mathbf{x})\Big)^{\mathrm{T}}. \tag{4.99}$$

We form a matrix using $\mathbf{w}(\mathbf{x})$ as

$$\mathbf{W} = \left[\mathbf{w}(\mathbf{x}_1)\ \mathbf{w}(\mathbf{x}_2)\ \mathbf{w}(\mathbf{x}_3) \ldots \mathbf{w}(\mathbf{x}_M)\right], \tag{4.100}$$

for some $\{\mathbf{x}_i\}_{i=1}^{M}$.

Regularity condition. *There exists vectors* $\{\mathbf{x}_i\}_{i=1}^{M} \neq \mathbf{0}$*, where* $M > (2(K-1)+1)$ *and constants* $\{p_i\}_{i=1}^{M} > 0$ *such that*

$$\mathbf{rank}\Big(\mathbf{W}\Big) = (2(K-1)+1), \tag{4.101}$$

$$\sum_{i=1}^{M} p_i \mathbf{w}(\mathbf{x}_i) = \mathbf{0}_{(2(K-1)+1)}. \tag{4.102}$$

***Remark* 4.4.1.** *The regularity condition implies that there does not exist any hyperplane passing through the origin such that all points* $\{\mathbf{w}(\mathbf{x}_i)\}_{i=1}^{M}$ *lie on one side of the hyperplane. This is seen as follows: For any non-zero* $\tilde{\mathbf{a}} \in \mathcal{R}^{(2(K-1)+1)}$*, taking the inner product w.r.t.* $\tilde{\mathbf{a}}$ *on both sides of* (4.102)*, we have* $\sum_{i=1}^{M} p_i(\tilde{\mathbf{a}}^{\mathrm{T}}\mathbf{w}(\mathbf{x}_i)) = 0$ *which implies that* $\tilde{\mathbf{a}}^{\mathrm{T}}\mathbf{w}(\mathbf{x}_i) \geq 0$ *or* $\tilde{\mathbf{a}}^{\mathrm{T}}\mathbf{w}(\mathbf{x}_i) \leq 0$ *for all* $i = 1, 2, \ldots, M$ *is not possible since* $\{p_i\}_{i=1}^{M} > 0$*. The special case of* $\tilde{\mathbf{a}}^{\mathrm{T}}\mathbf{w}(\mathbf{x}_i) = 0$ *for all* $i = 1, 2, \ldots, M$ *implies* $\mathbf{rank}\,(\mathbf{W}) <$

$(2(K-1)+1)$ *which contradicts with* (4.101)*. Hence, for any* $\tilde{\mathbf{a}}$*, we must have*

$$\tilde{\mathbf{a}}^{\mathrm{T}}\mathbf{w}(\mathbf{x}_i) < 0,\ \tilde{\mathbf{a}}^{\mathrm{T}}\mathbf{w}(\mathbf{x}_j) > 0 \text{ for some } i \text{ and } j, i \neq j. \tag{4.103}$$

The above regularity condition essentially requires a certain structure for our set $\mathcal{Y}$. This condition is necessary to prove the losslessness of the S-procedure. We now have the following theorem:

Theorem 4.4.1. *Assume* $\mathcal{Y}$ *satisfies the regularity condition. Denote* $\mathrm{cov}\,(X)$ *as the convex hull (smallest convex set enclosing* X*) of the set* X*. If* $\mathcal{Y} \cap \mathcal{N} = \emptyset$ *implies* $\mathrm{cov}\,(\mathcal{Y}) \cap \mathcal{N} = \emptyset$ *then the S-procedure is lossless.*[1]

Proof. See Publication V. □

The regularity condition is necessary for the S-procedure to be lossless. In Publication V, it is shown that $\mathcal{Y}$ indeed satisfies the regularity condition. Thus, Theorem 4.4.1 can be applied to the set $\mathcal{Y}$. Furthermore, for the S-procedure to be lossless, Theorem 4.4.1 states that $\mathcal{Y} \cap \mathcal{N} = \emptyset$ must imply $\mathrm{cov}\,(\mathcal{Y}) \cap \mathcal{N} = \emptyset$. We now have the following lemma:

Lemma 4.4.1. $\mathcal{Y} \cap \mathcal{N} = \emptyset \implies \mathrm{cov}\,(\mathcal{Y}) \cap \mathcal{N} = \emptyset$.

Proof. See Publication V. □

With these facts in place, we are now ready to prove strong duality between $(\mathcal{S})$ and $(\mathcal{DS})$.

***Remark* 4.4.2.** *Since* $\mathbf{x} \in \mathcal{C}^{N+1}$ *and* $\underline{\boldsymbol{\delta}} \in \mathcal{C}^{N}$*, we have that* $\Pi \subseteq \mathcal{Y}$.

Proposition 4.4.1. *Strong duality holds between* $(\mathcal{S})$ *and* $(\mathcal{DS})$ *and the optimal value is attained.*

Proof. The primal problem $(\mathcal{S})$ and its dual $(\mathcal{DS})$ are given by (4.89) and (4.90) respectively. From Remark 4.4.2, we have $\Pi \subseteq \mathcal{Y}$ and after using Lemma 4.4.1 and Theorem 4.4.1, we have that $\Pi \cap \mathcal{N} = \emptyset$ is equivalent to the matrix inequality in (4.90), i.e., the primal and dual problem are the same. □

Solving $(\mathcal{S})$***:*** Proposition 4.4.1 states that the optimal value of $(\mathcal{S})$ and $(\mathcal{DS})$ are the same, and we would like to find a $\underline{\boldsymbol{\delta}}$ in the constraint set that

[1]The result is not just limited to the quadratic forms that describe $\mathcal{Y}$, and it is applicable in general to any quadratic form.

attains this value. At this point $\underline{\boldsymbol{\delta}}_{\rm pc}$, the first-order Karush-Kuhn-Tucker conditions must be satisfied, i.e., there exists multipliers α_l^* and β_l^* s.t

$$\left(\tilde{\mathbf{M}} + \sum_{l=1}^{K-1} \alpha_l^* \tilde{\mathbf{P}}_l^R + \beta_l^* \tilde{\mathbf{P}}_l^I\right) \underline{\boldsymbol{\delta}}_{\rm pc} = \lambda^* \underline{\boldsymbol{\delta}}_{\rm pc}. \tag{4.104}$$

The above equation indicates that the minimizer of $(\mathcal{S})$ is an eigenvector of $\left(\tilde{\mathbf{M}} + \sum_{l=1}^{K-1} \alpha_l^* \tilde{\mathbf{P}}_l^R + \beta_l^* \tilde{\mathbf{P}}_l^I\right)$ with the eigenvalue equal to the optimal value of λ^*. Thus, the goal is to find λ^*, α_l^* and β_l^* by solving $(\mathcal{DS})$ using which we find $\underline{\boldsymbol{\delta}}_{\rm pc}$ from (4.104). The multipliers are obtained by forming the dual to $(\mathcal{DS})$ and solving both in tandem. The dual to $(\mathcal{DS})$ is given by [135]

$$\begin{aligned}(\mathcal{DDS}) : \text{Minimize } & \text{TR}\left(\tilde{\mathbf{M}}\mathbf{Z}\right) \text{ s.t } \text{TR}\left(\mathbf{Z}\right) = 1, \\ & \text{TR}\left(\tilde{\mathbf{P}}_l^R \mathbf{Z}\right) = 0,\ \text{TR}\left(\tilde{\mathbf{P}}_l^I \mathbf{Z}\right) = 0,\ \mathbf{Z} \succeq 0,\end{aligned} \tag{4.105}$$

where the matrix $\mathbf{Z}$ is the optimization variable and $\text{TR}(\cdot)$ denotes trace of a matrix. The problem $(\mathcal{DS})$ and its dual $(\mathcal{DDS})$ are semidefinite programs (SDP) which are nonlinear convex programs over the cone of positive semidefinite matrices [135]. In general, semidefinite programming solves for the multipliers in $(\mathcal{DS})$ and the matrix $\mathbf{Z}$ in $(\mathcal{DDS})$ simultaneously using interior-point methods specific for semidefinite programming. For a classic treatment on these methods, see [131, 135, 136].

Suboptimal schemes

The PNC scheme delivers an estimate of $\boldsymbol{\delta}$ using (4.104), where the multipliers λ^*, α_l^* and β_l^* are obtained by solving the semidefinite program of (4.105) [135]. In general, the computational complexity in solving an SDP can be high especially if N is large [136]. In this section, we seek optimization problems that have lesser complexity while at the same time ensuring that estimates satisfy the phase noise geometry. A suboptimal solution can be obtained by recognizing that the phase noise spectral geometry of (4.77) manifests itself as constant-magnitude time-domain samples. Such an approach is shown in Fig. 4.5, where the optimization problem $(\mathcal{AH})$ is given by

$$(\mathcal{AH}) : \quad \text{Minimize } \ \mathcal{L}(\underline{\boldsymbol{\delta}}) = \underline{\boldsymbol{\delta}}^\dagger \tilde{\mathbf{M}} \underline{\boldsymbol{\delta}} \ \text{ s.t } \ \underline{\boldsymbol{\delta}} \in \Omega, \tag{4.106}$$

where Ω is some constraint set that $\underline{\boldsymbol{\delta}}$ must belong to. After obtaining the solution to $(\mathcal{AH})$, we apply an inverse N-point DFT operation, normalize the time-domain samples to have unit-magnitude and finally get back to the frequency domain to obtain our modified solution which satisfies the

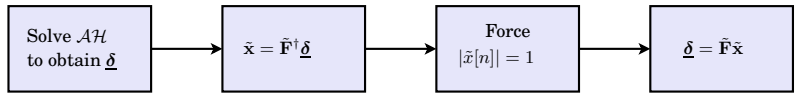


Figure 4.5. A suboptimal approach to force $\underline{\boldsymbol{\delta}}$ to satisfy (4.77). Then $\boldsymbol{\delta} = \mathbf{T}\underline{\boldsymbol{\delta}}$ will also satisfy (4.73) when $\mathbf{T}$ is chosen as a PPT. If $\mathbf{T}$ is not chosen as a PPT then we forcefully normalize the magnitude of the time-domain samples of $\mathbf{F}^{\dagger}\boldsymbol{\delta}$.

phase noise geometry. The added computational complexity is two DFT operations and another set of linear operations to forcefully normalize the magnitude of the time-domain samples.

The advantage of the suboptimal approach in Fig. 4.5 is that it allows us to choose a computationally less intensive optimization problem by proper choice of the constraint set Ω in $\mathcal{AH}$, while at the same time ensure that the resulting estimate satisfies the phase noise geometry at the cost of slightly increased complexity. In the following, we consider possible constraint sets for Ω other than the phase noise constraints, and after obtaining our estimate $\underline{\boldsymbol{\delta}}$, we perform the operations in Fig. 4.5.

Unit-norm Constraint (UNC): Rather than impose all the constraints in (4.77), we consider only the unit-norm constraint. Specifically, the optimization problem is given by

$$(\mathcal{U}): \quad \text{Minimize} \;\; \mathcal{L}(\underline{\boldsymbol{\delta}}) = \underline{\boldsymbol{\delta}}^{\dagger}\tilde{\mathbf{M}}\underline{\boldsymbol{\delta}} \;\; \text{s.t} \;\; \underline{\boldsymbol{\delta}}^{\dagger}\underline{\boldsymbol{\delta}} = 1. \tag{4.107}$$

The minimum value of the above problem is the minimum eigenvalue of $\tilde{\mathbf{M}}$ with the associated eigenvectors as the minimizers. If the eigenvalues are distinct then the minimizer is unique.

Linear Constraint (LC): In general, linear constraints are beneficial because they yield computationally less intensive optimization problems compared to second and higher-order constraints. In [119], the authors make use of a linear constraint when minimizing the cost function $\mathcal{L}(\underline{\boldsymbol{\delta}}) = \underline{\boldsymbol{\delta}}^{\dagger}\left(\mathbf{L}^{\dagger}\tilde{\mathbf{M}}\mathbf{L}\right)\underline{\boldsymbol{\delta}}$, where $\mathbf{L}$ is given in (4.74). We can generalize the problem by choosing different transformation matrices. Specifically, the problem is given by

$$(\mathcal{P}): \text{Minimize}\; \mathcal{L}(\underline{\boldsymbol{\delta}}) = \underline{\boldsymbol{\delta}}^{\dagger}\tilde{\mathbf{M}}\underline{\boldsymbol{\delta}} \;\text{s.t}\; \frac{1}{2}\left(\underline{\boldsymbol{\delta}}^{\dagger}\tilde{\mathbf{e}}_1 + \tilde{\mathbf{e}}_1^{\dagger}\underline{\boldsymbol{\delta}}\right) = 1, \tag{4.108}$$

where $\tilde{\mathbf{M}} = \mathbf{T}^{\dagger}\mathbf{M}\mathbf{T}$ and $\tilde{\mathbf{e}}_1 = [1\; 0, \ldots, 0]^{\mathrm{T}}$ is an $N \times 1$ column vector. After solving $(\mathcal{P})$, we must perform the operations in Fig. 4.5 so that the resulting vector satisfies the phase noise geometry. These operations are not done in [119]. It is easily seen that $(\mathcal{P})$ is a convex program. In fact,

Table 4.1. Phase noise estimation complexity for each iteration of the decision-directed scheme in Table 4.2.

Constraint used	Computations-per-iteration
PNC	$O\left(N^{4.5}\right)$
UNC	$O\left(N^3\right) + O\left(N\log(N)\right)$
LC	$O\left(N^3\right) + O\left(N\log(N)\right)$

for such a problem, a closed-form solution exists. At the optimal solution $\underline{\boldsymbol{\delta}}_{\mathrm{lc}}$, the first-order Karush-Kuhn-Tucker condition must be satisfied and is given by

$$\tilde{\mathbf{M}}\underline{\boldsymbol{\delta}}_{\mathrm{lc}} + \frac{\lambda}{2}\tilde{\mathbf{e}}_1 = 0 \;\text{ implying }\; \underline{\boldsymbol{\delta}}_{\mathrm{lc}} = -\frac{\lambda}{2}\tilde{\mathbf{M}}^{-1}\tilde{\mathbf{e}}_1. \tag{4.109}$$

The Lagrange multiplier λ is obtained by substituting $\underline{\boldsymbol{\delta}}_{\mathrm{lc}}$ in the linear constraint and solving to obtain $\lambda = -\frac{2}{\tilde{\mathbf{e}}_1^{\dagger}\tilde{\mathbf{M}}^{-1}\tilde{\mathbf{e}}_1}$.

Computational Complexity of PNC, UNC and LC

In Publication V, the computational complexity of PNC, UNC and LC are studied. The overall, general case complexity for these schemes is shown in Table 4.1. From Table 4.1, we see that PNC has the highest computational complexity. This arises mainly because a semidefinite program needs to be solved to arrive at a phase noise estimate using the PNC scheme. In general, although convex in nature, semidefinite programs have a higher computational complexity and are solved using interior-point methods. See, for example, [131, 135, 136], where these methods and their complexity are described in detail.

For the LC scheme, we need to solve (4.109) which is nothing but solving a system of linear equations and, in general, the complexity is $O(N^3)$. The added complexity of $O\left(N\log(N)\right)$ comes from performing the operations in Fig. 4.5. The main computational complexity considered in Fig. 4.5 are the DFT and IDFT operations. This is assuming that $\mathbf{T}$ is chosen as PPT. If set to some other transformation matrix then the complexity is $O\left(N_{\mathrm{c}}\log(N_{\mathrm{c}})\right)$.

For the UNC scheme, the eigenvector associated with the minimum eigenvalue of the matrix $\tilde{\mathbf{M}}$ needs to be determined. In general, this requires $O(N^3)$ operations, and after performing the operations in Fig. 4.5, we have the total amount as shown in Table 4.1. See, for example, [137] for an excellent treatise on eigenvalue algorithms.

Table 4.2. Decision-directed scheme.

Initialization step
(i) Choose a particular transformation matrix $\mathbf{T}$.
(ii) Estimate channel: $\hat{H}_p = \frac{r_p}{s_p}$, where p is pilot subcarrier index.
(iii) Interpolate $\hat{H}_p$ to obtain $\hat{H}_j,\ j = 0, 1, \ldots, N_\mathrm{c} - 1$.
(iv) Set $z_j = \frac{r_j}{\hat{H}_j},\ j = 0, 1, \ldots, N_\mathrm{c} - 1$.
(v) $\hat{\mathbf{s}} = \mathrm{Decode}\,(\mathbf{z})$.
(vi) Form $\mathbf{S} = \mathrm{diag}(\hat{\mathbf{s}})$, $\mathbf{B} = \mathbf{F}_\mathrm{t}^\dagger \mathbf{S}^\dagger \mathbf{S} \mathbf{F}_\mathrm{t}$ and $\mathbf{P}_\mathrm{r} = \mathbf{S}\mathbf{F}_\mathrm{t}\mathbf{B}^{-1}\mathbf{F}_\mathrm{t}^\dagger\mathbf{S}^\dagger$.
(vii) Form $\mathbf{M} = \mathbf{R}^\mathrm{T}\mathbf{R}^* - \mathbf{R}^\mathrm{T}\mathbf{P}_\mathrm{r}^\mathrm{T}\mathbf{R}^*$ and then $\tilde{\mathbf{M}} = \mathbf{T}^\dagger\mathbf{M}\mathbf{T}$.
(viii) Set *loop* variable to number of iterations.
Decision-feedback step
for $i = 1$ to *loop*
Phase noise estimation step:
(i) Solve optimization problems in Section 4.4.3 to obtain $\underline{\boldsymbol{\delta}}$.
(ii) Set $\boldsymbol{\delta} = \mathbf{T}\underline{\boldsymbol{\delta}}$.
Channel estimation step:
(i) Form the $\mathbf{V}$ matrix using $\boldsymbol{\delta}$.
(ii) $\hat{\mathbf{h}}_\mathrm{ML} = \left(\mathbf{F}_\mathrm{t}^\dagger\mathbf{S}^\dagger\mathbf{S}\mathbf{F}_\mathrm{t}\right)^{-1}\mathbf{F}_\mathrm{t}^\dagger\mathbf{S}^\dagger\mathbf{V}^\dagger\mathbf{r}$.
Phase noise compensation step
(i) Remove phase noise by performing $\mathbf{z} = \mathbf{V}^\dagger\mathbf{r}$.
Channel equalization step
(i) Zero-forcing equalization: $\mathbf{y} = \left(\mathrm{diag}\left(\mathbf{F}_\mathrm{t}\hat{\mathbf{h}}_\mathrm{ML}\right)\right)^{-1}\mathbf{z}$.
Symbol detection step
(i) $\hat{\mathbf{s}} = \mathrm{Decode}\,(\mathbf{y})$.
(ii) Form $\mathbf{S} = \mathrm{diag}\,(\hat{\mathbf{s}})$, $\mathbf{B} = \mathbf{F}_\mathrm{t}^\dagger\mathbf{S}^\dagger\mathbf{S}\mathbf{F}_\mathrm{t}$ and $\mathbf{P}_\mathrm{r} = \mathbf{S}\mathbf{F}_\mathrm{t}\mathbf{B}^{-1}\mathbf{F}_\mathrm{t}^\dagger\mathbf{S}^\dagger$.
(iii) Form $\mathbf{M} = \mathbf{R}^\mathrm{T}\mathbf{R}^* - \mathbf{R}^\mathrm{T}\mathbf{P}_\mathrm{r}^\mathrm{T}\mathbf{R}^*$ and then $\tilde{\mathbf{M}} = \mathbf{T}^\dagger\mathbf{M}\mathbf{T}$.
end

Numerical Results

Let us now present some numerical results on the geometry-based phase noise minimization schemes originally reported in Publication V. In Publication V, a *decision-directed* scheme is used for estimating phase noise and channel. Such a scheme is shown in Table 4.2. In Table 4.2, at the symbol detection step, the operation $\mathrm{Decode}\,(\cdot)$ includes the steps of symbol constellation demodulation and channel decoding. Estimates of s_j are obtained by performing sequentially: phase noise and channel estimation;

phase noise removal or compensation; channel equalization; and, finally, symbol detection. The phase noise and channel estimation step requires knowledge of the S matrix which comprises of the transmitted symbols s_j, which we do not know. Instead estimates of s_j can be used to form S as seen in Table 4.2. The obtained phase noise and channel estimates are then used to obtain new estimates of s_j. These new estimates are then used to improve the phase noise and channel estimates, thereby, resulting in a decision-directed feedback scheme.

We now compare the impact of phase noise estimation schemes on the decision-directed scheme of Table 4.2. The phase noise estimation step is implemented using PNC, UNC and LC schemes. The simulations are run with the following system parameters: The number of subcarriers $N_{\mathrm{c}} = 512$, subcarrier spacing $f_{\mathrm{sub}} = 15$ kHz and bandwidth is 7.7 MHz. Phase noise is modeled as a Wiener process with the oscillator PSD 3-dB bandwidth $f_{3\mathrm{dB}} = 300$ Hz. The length of the vector $\underline{\delta}$ is set to a value of seven, i.e., $N = 7$. The percentage of scattered pilots is set to 8% and symbol constellation is 16-QAM. The channel is Rayleigh fading with 10 exponentially decaying taps, and coherence bandwidth is set to 800 kHz. A $1/2$-rate convolutional encoder $[133, 171]$ with constraint length of 7 is used for channel encoding. For channel decoding, a soft-decision Viterbi decoder of decoding depth equal to five times the constraint length is used.

Figure 4.6 shows the coded BER plots of the decision-directed scheme of Table 4.2. The curve corresponding to the case of perfect phase noise and channel knowledge serves as a benchmark in judging the performance of the phase noise estimation schemes. From the figure, PNC offers the lowest coded BER for the decision-directed scheme with UNC performing the second best and LC performing the worst. Although PNC provides for the lowest BER compared to others, its computational complexity can be quite high. Thus, depending upon the system design requirement, UNC and LC maybe better alternatives from the point of view of computational complexity. See Section 4.4.3 for a discussion on the complexity. Figure 4.7 demonstrates the reduction in the coded BER as the number of iterations in the decision-directed scheme are increased. This reduction in BER justifies the efficacy of using a decision-directed scheme from the point of view of performance. The decision-directed scheme will yield a low BER provided that most of the initial estimates of the transmitted symbols are correct. For such a situation to happen, the phase noise process must be slow-varying in nature otherwise the initial estimates will be poor, and

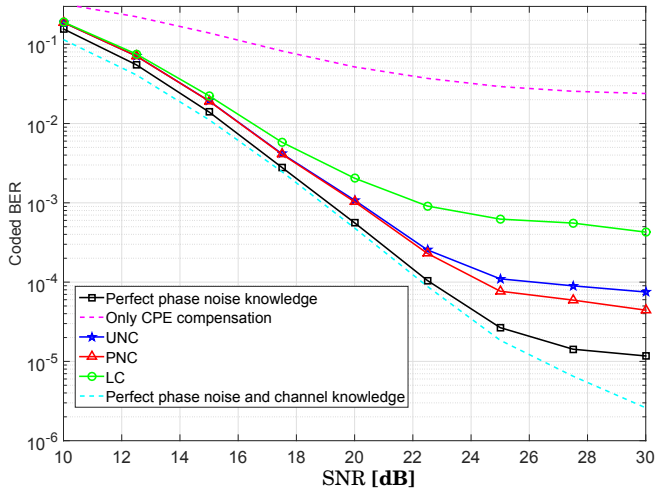


Figure 4.6. Comparison of average coded BER vs. SNR for the proposed schemes. The number of iterations in the decision-directed scheme is set to a value of 5. The matrix $\mathbf{T}$ used is the piecewise constant PPT of (4.80).

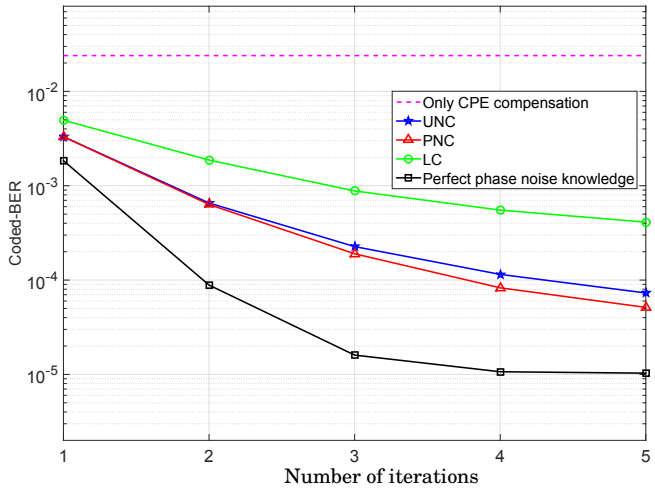


Figure 4.7. Coded BER as a function of the number of iterations in the decision-directed scheme. SNR is set to 30 dB, $f_{3\text{dB}} = 300$ Hz, $f_{\text{sub}} = 15$ kHz and, hence, $\rho = \frac{f_{3\text{dB}}}{f_{\text{sub}}} = 0.02$. The matrix $\mathbf{T}$ is set as the PC-PPT of (4.80).

the decision-directed scheme will not converge to a lower BER. For the simulations, the value of $\rho = 0.02$ which, from a system design point of view, is quite high. Even for this value of ρ, the decision-directed scheme converges well.

4.5 Discussion

This chapter presents an overview and classification of some of the state-of-the-art phase noise estimation algorithms for OFDM. By no means is this overview comprehensive or complete, and the reader is encouraged to see the references cited in this chapter for a more complete picture. In addition, this chapter also presents the two novel phase noise estimation algorithms of Publications V and VI, respectively.

Estimation schemes in OFDM, in the context of phase noise, can be classified into three types: Isolated estimation; Semi-joint estimation; and Fully-joint estimation. The goal here is to estimate channel, phase noise and the transmitted symbols. In isolated approaches, these three quantities are estimated independently of each other while assuming knowledge of the other. For example, for phase noise estimation, it is assumed that a channel estimate and symbols estimate is readily available which is used in estimating the phase noise. The symbol estimates are either tentative decisions on transmitted symbols or are the pilot symbols themselves. Most phase noise estimation algorithms initially developed fall into the isolated category, where channel estimates were obtained using methods assuming no phase noise in the system. For slow-varying phase noise processes, such an approach is acceptable, however, for fast-varying processes, the performance will be inferior.

Poor channel estimates resulting from ignoring phase noise results in poor symbol estimates, thereby, yielding a poor BER performance. This recognition has led to *joint channel and phase noise* estimation algorithms, where the primary focus is to obtain reliable channel estimates while taking the effect of phase noise into account. To do so, most algorithms make use of the pilot phase of transmission, where all the transmitted symbols are known a-priori at the receiver side. In applications with no dedicated pilot phase transmission, tentative decisions on the transmitted symbols can be used. A semi-joint estimation scheme would use the channel estimate obtained using the joint approach to estimate the phase noise and transmitted symbols during the data phase of transmission, wherein the phase noise and symbol estimation steps are performed independently.

To obtain statistically optimal channel, phase noise and symbol estimates, a fully-joint approach must be used; During the pilot phase, a joint channel and phase noise estimation algorithm is used; and during the data phase, a joint phase noise and symbol estimation algorithm is used.

Joint approaches, in general, are computationally more complex simply because a larger number of parameters are estimated. A fully-joint approach will incur delay into the system which may not be tolerable in low-latency wireless applications. In contrast, isolated approaches may be preferred, where separate and computationally less intensive channel, phase noise and symbol estimation algorithms are used such that overall complexity is less than when using a fully-joint approach.

The work in Publication VI contributes to the area of joint channel and phase noise estimation, where the goal is to obtain reliable channel estimates. It improves the method, originally proposed in [119], by choosing subspaces that always contain the desired phase noise vector. The phase noise estimate obtained in [119] is reasonable only for slow-varying phase noise processes, however, the method breaks down for fast-varying phase noise processes. On the other hand, the work in Publication VI utilizes subspace information that does not depend upon the nature of the phase noise process. Such a scheme is, hence, also useful at moderate to fast-varying phase noise processes as well as slow-varying ones.

The work in Publication V can be used in an isolated approach or a joint approach where the objective is to obtain optimal channel estimates in the presence of phase noise. The phase noise estimation step utilizes information on the geometry of the spectral components of the complex exponential of the phase noise process. Such an estimate is obtained by solving a quadratic cost function subject to quadratic equality constraints that involve permutation matrices. The complexity of solving such an algorithm can be high depending upon the dimensionality of the estimated phase noise vector. In order to reduce the complexity, the work in Publication V also proposes suboptimal schemes that achieve the same objective of satisfying the phase noise geometry.

5. Conclusion

OFDM has become a popular choice for modulation and is widely used in many communication systems, especially in wireless systems. For example, in the latest fourth-generation of wireless systems, it is used in the downlink. It is also a contender for the upcoming fifth-generation wireless systems as well. It facilitates simple implementation of the baseband modulator and demodulator and, most importantly, it trivializes the task of channel equalization. It, however, has its drawbacks and, by far, the most important being its sensitivity to RF-impairments. This thesis contributes to the field of analysis and estimation for OFDM systems impaired by phase noise which forms one type of RF-impairment.

5.1 Contributions in OFDM under Oscillator Phase Noise

With regard to performance analysis, this thesis contributes by providing new closed-form analytical expressions of capacity for OFDM systems impaired by phase noise. The capacity analysis is also extended to include the effect of carrier frequency offset. The resulting expressions provide quantitative as well as qualitative insight on the relationship between the phase noise process and capacity. Through these expressions, a clear degradation in capacity of the OFDM system in the presence of phase noise is seen. Fortunately, this degradation can be controlled either by proper choice of oscillator design, or by adjusting the OFDM system parameters, or by performing phase noise estimation and compensation.

This thesis also makes two new contributions to the field of phase noise estimation in OFDM. Specifically, two novel aspects about the desired phase noise parameter are used during the estimation step. In the first contribution, subspace-based information is used, i.e., possible subspaces in which the desired phase noise spectral vector may lie are explored. This

subspace-based approach avoids the need to make assumptions about the nature of the phase noise process which is generally used in the literature. For example, phase noise processes are assumed to be slow-varying in nature which is, no doubt, a reasonable assumption. For moderate or fast-varying phase noise processes, the proposed subspace-based approach will perform better compared to algorithms specifically designed with the slow-varying assumption.

In the second contribution, information on the geometry of the desired phase noise spectral vector is used in the estimation step. The goal is to estimate the spectral vector of the complex exponential of the phase noise process. It is shown in this thesis that this spectral vector is always drawn from a non-convex set which can be expressed using a set of quadratic forms that involve permutation matrices. This geometry is nothing but a frequency domain manifestation of the *constant magnitude* property of the complex exponential function. The constant magnitude property is a well-known fact but its equivalent frequency domain manifestation has not been observed and utilized in the research community.

5.2 Contributions in Applied Statistics and Optimization Theory

This thesis also presents some new fundamental results in the fields of applied statistics and optimization theory. These results are application independent and can be applied wherever suitable. As an example, pertaining to the field of statistics, the PDF of a sum of correlated gamma random variables with a normalized covariance matrix of any rank is derived. The state-of-the-art result was limited to the full-rank case as derived in [96]. The framework, based on the work done by Moschopoulos in [95], naturally allows to extend the result to deriving the PDF of a sum of correlated gamma and Gaussian distributed random variables.

This thesis also contributes to the field of optimization theory and provides some new results on the losslessness of the S-procedure that involve equality constraints. The S-procedure is a method of replacing a set of quadratic equalities or inequalities with a linear-matrix-inequality. Conditions for the S-procedure to be lossless for the case of quadratic inequalities is well-established and used extensively. This thesis fills the void by providing conditions for the S-procedure to be lossless for any number of quadratic equality constraints.

5.3 Directions of Future Work

We end this chapter with a brief treatise on possible research directions related to OFDM and phase noise. The next-generation of wireless systems, namely 5G, have set forth gigantic goals on data rate, latency and spectral efficiency to name a few. For example, peak data rates are expected to deliver three orders of magnitude more than the current 4G systems [138]. The key physical layer technologies that promise to deliver such massive data rates are primarily: *massive MIMO*, *millimeter wave communications* and *heterogeneous networks* [4]. In the context of phase noise, there is abundant research that shows a performance degradation for MIMO systems corrupted by phase noise, for example, when performing beamforming [139]. The effect of phase noise on MIMO systems serves as an indicator of what to expect with regard to massive MIMO systems that employ hundreds of antennas at the base station. This indeed marks the beginning of a new research area dedicated to analyzing and addressing the effects of phase noise and RF-impairments, in general, on massive MIMO systems. Some new studies on the topic can be found in [140] and [141].

An important aspect to delivering high data rates is the notion of spectral efficiency which is intrinsically linked to the underlying waveform and symbol constellation. For example, in LTE, the OFDM waveform is used. It is a popular contender for the upcoming 5G systems, however, alternatives are being sought after [142]. Two main reasons for seeking alternatives are: low spectral efficiency of OFDM; and stringent synchronization requirements [143]. By synchronization, we refer to timing and frequency synchronization. For example, in a multi-user uplink scenarios, the base station needs to estimate the timing and frequency offset of all the users it services in order to avoid inter-block and multi-user interference. This stringent requirement arises fundamentally because of the susceptibility of OFDM to timing and frequency offset. Frequency offset can be viewed as a deterministic version of phase noise. Alternative multi-carrier waveforms, such as the filter bank multi-carrier, are more robust to synchronization errors compared to OFDM and also have higher spectral efficiency [144]. The impact of phase noise on these waveforms is not well-known, and we envisage a new research field dedicated to this area.

A. Appendix

A.1 Proof of Theorem 3.4.1

In order to derive the PDF, we first form the moment generating function (MGF) of Y and simplify the resulting expression. We then obtain the PDF after applying an inverse Laplace transform. The MGF of Y can be obtained on similar lines as done in Publication I to obtain

$$M_Y(s) = \int_{-\infty}^{\infty} \frac{e^{sd} e^{s(\mathbf{Pa})^{\mathrm{T}}\mathbf{y}} e^{-\frac{1}{2}\mathbf{y}^{\mathrm{T}}[\mathbf{\Delta}^{-1} - s2\mathbf{PBP}^{\mathrm{T}}]\mathbf{y}} \mathbf{dy}}{(2\pi)^{\frac{R}{2}} |\mathbf{\Delta}|^{\frac{1}{2}}}, \tag{1.1}$$

where $\mathbf{P} = \begin{bmatrix}\mathbf{I}_{R\times R} & \mathbf{0}_{R\times(N-R)}\end{bmatrix}\mathbf{C}^{\mathrm{T}}$. The matrix $\mathbf{C}$ is obtained from $\mathbf{M_x} = \mathbf{C\Sigma C}^{\mathrm{T}}$ which is an eigendecomposition of $\mathbf{M_x}$. The matrix $\mathbf{\Delta}$ is diagonal whose elements are the non-zero eigenvalues of $\mathbf{M_x}$. The integral in (1.1) can be solved by using the result from [145] to obtain

$$\begin{aligned} M_Y(s) &= e^{sd} \prod_{i=1}^{R} \frac{e^{\frac{1}{2}\left[s(\mathbf{V}^{\mathrm{T}}\mathbf{\Delta}^{1/2}\mathbf{Pa})^{\mathrm{T}}(\mathbf{I}-s\mathbf{\Lambda})^{-1}(\mathbf{V}^{\mathrm{T}}\mathbf{\Delta}^{1/2}\mathbf{Pa})s\right]}}{(1-s\lambda_i)^{\frac{1}{2}}} \\ &= e^{sd} \prod_{i=1}^{R} \frac{e^{\frac{s^2 c_i^2/2}{1-s\lambda_i}}}{(1-s\lambda_i)^{\frac{1}{2}}} \qquad (1.2) \\ &= K_c e^{sd} e^{-s\tau} \prod_{i=1}^{R} \frac{e^{\frac{b_i/2}{1-s\lambda_i}}}{(1-s\lambda_i)^{\frac{1}{2}}}, \qquad (1.3) \end{aligned}$$

where $\mathbf{V\Lambda V}^{\mathrm{T}}$ is the eigenvalue decomposition of $2\Delta^{1/2}\mathbf{PBP}^{\mathrm{T}}\Delta^{1/2}$ with $\{\lambda_i\}_{i=1}^{R}$ being the ordered non-zero eigenvalues (λ_1 is the minimum). The vector $\mathbf{c} = \left(\mathbf{V}^{\mathrm{T}}\Delta^{1/2}\mathbf{Pa}\right)$ whose elements are denoted by c_i. The terms $b_i = c_i^2/\lambda_i^2$, $\tau = \frac{1}{2}\sum_{i=1}^{R} b_i\lambda_i$ and $K_c = e^{-\frac{1}{2}\sum_{i=1}^{R} b_i}$. The final step of (1.3) is obtained by applying partial fraction expansion on the exponential term in (1.2).

We now use an original idea proposed in [95] that equivalently expresses

the product term in (1.3) as a summation term. First, we rewrite

$$1 - s\lambda_i = (1 - s\lambda_1)\,\frac{\lambda_i}{\lambda_1}\left[1 - \left(\frac{1 - \lambda_1/\lambda_i}{1 - s\lambda_1}\right)\right]. \tag{1.4}$$

Taking the natural logarithm on both sides of (1.3) and, after using (1.4), we have

$$\ln\left(M_Y(s)\right) = \ln\left(K_{\rm c} e^{sd} e^{-s\tau}\right) + \ln\left(\left(1 - s\lambda_1\right)^{-R/2} K_{\rm p}\right) + \frac{1}{2}\sum_{i=1}^{R} \frac{b_i\left[1 - \left(\frac{1-\lambda_1/\lambda_i}{1-s\lambda_1}\right)\right]^{-1}}{(1 - s\lambda_1)\,\frac{\lambda_i}{\lambda_1}} - \frac{1}{2}\sum_{i=1}^{R} \ln\left(1 - \left(\frac{1 - \lambda_1/\lambda_i}{1 - s\lambda_1}\right)\right), \tag{1.5}$$

where $K_{\rm p} = \prod_{i=1}^{R}\left(\frac{\lambda_1}{\lambda_i}\right)^{\frac{1}{2}}$. Using the expansions $\ln(1+x) = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{x^k}{k}$ and $(1+x)^{-1} = \sum_{k=1}^{\infty}(-1)^{k-1}x^{k-1}$, (1.5) simplifies to [97]

$$\begin{aligned}
&\ln\left(M_Y(s)\right) = \ln\left(K_c e^{sd} e^{-s\tau}\right) + \ln\left(\left(1 - s\lambda_1\right)^{-R/2} K_p\right) + \\
&\frac{1}{2}\sum_{i=1}^{R} \frac{b_i}{(1 - s\lambda_1)\,\frac{\lambda_i}{\lambda_1}}\left(\sum_{k=1}^{\infty}\left(\frac{1 - \lambda_1/\lambda_i}{1 - s\lambda_1}\right)^{k-1}\right) + \frac{1}{2}\sum_{i=1}^{R}\left(\sum_{k=1}^{\infty}\left(\frac{1 - \lambda_1/\lambda_i}{1 - s\lambda_1}\right)^{k}\frac{1}{k}\right), \\
&= \ln\left(K_{\rm c} e^{sd} e^{-s\tau}\right) + \ln\left(\left(1 - s\lambda_1\right)^{-R/2} K_p\right) + \sum_{k=1}^{\infty} z_k \left(1 - s\lambda_1\right)^{-k},
\end{aligned} \tag{1.6}$$

where the coefficient z_k is given by the following equation:

$$z_k = \frac{1}{2}\sum_{i=1}^{R}\left(\frac{b_i\left(1 - \lambda_1/\lambda_i\right)^{k-1}}{\frac{\lambda_i}{\lambda_1}} + \frac{\left(1 - \lambda_1/\lambda_i\right)^{k}}{k}\right). \tag{1.7}$$

Taking the inverse logarithm on both sides of (1.6), we have

$$M_Y(s) = K e^{sd} e^{-s\tau}\left(1 - s\lambda_1\right)^{-R/2} e^{\sum_{k=1}^{\infty} z_k \left(1 - s\lambda_1\right)^{-k}}, \tag{1.8}$$

where $K = K_{\rm c}K_{\rm p}$. In [95], it is shown that the exponential term containing the summation series can be equivalently expressed using a summation series, i.e., $e^{\sum_{k=1}^{\infty} z_k \left(1 - s\lambda_1\right)^{-k}} = \sum_{k=0}^{\infty}\zeta_k\left(1 - s\lambda_1\right)^{-k}$, where the coefficients ζ_k and z_k are related by

$$\begin{aligned}
\zeta_0 = 1,\ \zeta_{k+1} &= \frac{1}{k+1}\sum_{i=1}^{k+1} i z_i \zeta_{k+1-i},\ k = 0, 1, \ldots \\
&= \frac{0.5}{k+1}\sum_{i=1}^{k+1}\left[\sum_{j=1}^{R}\left(1 - \lambda_1/\lambda_j\right)^{i}\left(1 + \frac{i b_j(\lambda_1/\lambda_j)}{(1 - \lambda_1/\lambda_j)}\right)\right]\zeta_{k+1-i}.
\end{aligned} \tag{1.9}$$

Using this equivalent representation, the MGF of Y is given by

$$M_Y(s) = K e^{sd} e^{-s\tau}\sum_{k=0}^{\infty}\frac{\zeta_k}{\left(1 - s\lambda_1\right)^{(R/2+k)}}, \tag{1.10}$$

and after taking the inverse Laplace transform of $M_Y(-s)$, we obtain the PDF of Y as given in (3.42).

A.2 Proof of Theorem 3.5.1

The theorem can be proved on similar lines as done in Section A.1 with minor modifications that take into account aspects of negativity of the eigenvalues as well as rank deficiency. Using the approach in Publication I, the MGF of $Y = \mathbf{x}^{\mathrm{T}}\left(\mathbf{a} + \mathbf{B}\mathbf{x}\right)$ is simplified to

$$M_Y(s) = \int_{-\infty}^{\infty} \frac{e^{s(\mathbf{P}\mathbf{a})^{\mathrm{T}}\mathbf{y}} e^{-\frac{1}{2}\mathbf{y}^{\mathrm{T}}[\mathbf{\Delta}^{-1} - s2\mathbf{P}\mathbf{B}\mathbf{P}^{\mathrm{T}}]\mathbf{y}}\mathbf{d}\mathbf{y}}{(2\pi)^{\frac{R}{2}}|\mathbf{\Delta}|^{\frac{1}{2}}}. \tag{1.11}$$

The integral in the above equation has a closed-form expression and the resulting expression is given by [145]

$$\begin{aligned} M_Y(s) &= \left(\prod_{i=1}^{R+L+Z}(1-s\lambda_i)^{-\frac{1}{2}}\right) e^{\frac{1}{2}\left[s(\mathbf{V}^{\mathrm{T}}\mathbf{\Delta}^{1/2}\mathbf{P}\mathbf{a})^{\mathrm{T}}(\mathbf{I}-s\mathbf{\Lambda})^{-1}(\mathbf{V}^{\mathrm{T}}\mathbf{\Delta}^{1/2}\mathbf{P}\mathbf{a})s\right]} \\ &= \left[\prod_{i\in\mathcal{R}} \frac{e^{\frac{1}{2}\frac{s^2c_i^2}{1-s\lambda_i}}}{(1-s\lambda_i)^{\frac{1}{2}}}\right]\left[\prod_{i\in\mathcal{L}} \frac{e^{\frac{1}{2}\frac{s^2c_i^2}{1-s\lambda_i}}}{(1-s\lambda_i)^{\frac{1}{2}}}\right] e^{\frac{1}{2}s^2\sum_{i\in\mathcal{Z}}c_i^2}, \end{aligned} \tag{1.12}$$

where λ_i denotes the eigenvalues of $2\Delta^{1/2}\mathbf{P}\mathbf{B}\mathbf{P}^{\mathrm{T}}\Delta^{1/2}$ with R number of positive eigenvalues, L number of negative eigenvalues and Z number of zero eigenvalues. The eigendecomposition of $2\Delta^{1/2}\mathbf{P}\mathbf{B}\mathbf{P}^{\mathrm{T}}\Delta^{1/2}$ is given by $\mathbf{V}\mathbf{\Lambda}\mathbf{V}^{\mathrm{T}}$. The index sets $\mathcal{R}$, $\mathcal{L}$ and $\mathcal{Z}$ correspond to the positive, negative and zero eigenvalues. The elements c_i are the components of the vector $\mathbf{c} = \mathbf{V}^{\mathrm{T}}\Delta^{1/2}\mathbf{P}\mathbf{a}$.

Applying partial fraction expansion on the exponents inside the paranthesis of (1.12), we have

$$M_Y(s) = K_c e^{-s\tau}\left[\prod_{i=1}^{R}\frac{e^{\frac{u_i/2}{1-s\beta_i}}}{(1-s\beta_i)^{\frac{1}{2}}}\right]\left[\prod_{i=1}^{L}\frac{e^{\frac{v_i/2}{1-s\gamma_i}}}{(1-s\gamma_i)^{\frac{1}{2}}}\right]\left[e^{\frac{1}{2}s^2\sum_{i\in\mathcal{Z}}c_i^2}\right], \tag{1.13}$$

where $u_i = (c_k^2/\lambda_k^2),\ k \in \mathcal{R}$, $v_i = (c_k^2/\lambda_k^2),\ k \in \mathcal{L}$, and the constant $K_c = e^{-\frac{1}{2}\left(\sum_{i=1}^{R}u_i + \sum_{i=1}^{L}v_i\right)}$. The delay parameter $\tau = \frac{1}{2}\left(\sum_{k\in\mathcal{R}}(\frac{c_k^2}{\lambda_k}) + \sum_{k\in\mathcal{L}}\left(\frac{c_k^2}{\lambda_k}\right)\right)$. In (1.13), we denote $\beta_i = \lambda_k,\ k \in \mathcal{R}$, where $i = 1, 2, \ldots, R$ and $\gamma_i = \lambda_k,\ k \in \mathcal{L}$, where $i = 1, 2, \ldots, L$. Equation (1.13) is a generalized version of (1.3) which only assumes $2\Delta^{1/2}\mathbf{P}\mathbf{B}\mathbf{P}^{\mathrm{T}}\Delta^{1/2}$ to be of full-rank with only positive eigenvalues. In contrast, the first parenthesis term represents the positive contribution of the eigenvalues, the second parenthesis term represents the negative contribution, and the third parenthesis results because of the rank deficiency. Note that the third parenthesis term represents the MGF of a Gaussian distributed random variable.

By using (1.4) and the approach following it, the first and second paran-

thesis terms can be effectively represented by

$$\left[\prod_{i=1}^{R}\frac{e^{\frac{u_i/2}{1-s\beta_i}}}{(1-s\beta_i)^{\frac{1}{2}}}\right]=\prod_{i=1}^{R}\left(\frac{\beta_1}{\beta_i}\right)^{\frac{1}{2}}\sum_{k=1}^{\infty}\frac{\eta_k}{(1-s\beta_1)^{(R/2+k)}}, \tag{1.14}$$

$$\left[\prod_{i=1}^{L}\frac{e^{\frac{v_i/2}{1-s\gamma_i}}}{(1-s\gamma_i)^{\frac{1}{2}}}\right]=\prod_{i=1}^{L}\left(\frac{\gamma_1}{\gamma_i}\right)^{\frac{1}{2}}\sum_{k=1}^{\infty}\frac{\zeta_k}{(1+s|\gamma_1|)^{(L/2+k)}}, \tag{1.15}$$

where β_1 and γ_1 are the minimum and maximum amoung β_i and γ_i, respectively and $|\gamma_1|$ denotes the absolute value of γ_1. The coefficients η_k and ζ_k are given by

$$\eta_0=1,\ \eta_k=\frac{0.5}{k+1}\sum_{i=1}^{k+1}\left[\sum_{j=1}^{R}(1-\beta_1/\beta_j)^i\left(1+\frac{iu_j}{((\beta_j/\beta_1)-1)}\right)\right]\eta_{k+1-i}, \tag{1.16}$$

$$\zeta_0=1,\ \zeta_k=\frac{0.5}{k+1}\sum_{i=1}^{k+1}\left[\sum_{j=1}^{L}(1-\gamma_1/\gamma_j)^i\left(1+\frac{iv_j}{((\gamma_j/\gamma_1)-1)}\right)\right]\zeta_{k+1-i}. \tag{1.17}$$

Taking the product of (1.14) and (1.15), we have

$$\left[\prod_{i=1}^{R}\frac{e^{\frac{u_i/2}{1-s\beta_i}}}{(1-s\beta_i)^{\frac{1}{2}}}\right]\left[\prod_{i=1}^{L}\frac{e^{\frac{v_i/2}{1-s\gamma_i}}}{(1-s\gamma_i)^{\frac{1}{2}}}\right]$$
$$=\prod_{i=1}^{R}\left(\frac{\beta_1}{\beta_i}\right)^{\frac{1}{2}}\prod_{i=1}^{L}\left(\frac{\gamma_1}{\gamma_i}\right)^{\frac{1}{2}}\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\frac{\eta_k\zeta_j}{(1-s\beta_1)^{(R/2+k)}\,(1+s|\gamma_1|)^{(L/2+k)}}, \tag{1.18}$$
$$=\prod_{i=1}^{R}\left(\frac{\beta_1}{\beta_i}\right)^{\frac{1}{2}}\prod_{i=1}^{L}\left(\frac{\gamma_1}{\gamma_i}\right)^{\frac{1}{2}}\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\eta_k\zeta_j\left(\sum_{l=1}^{R_k}\frac{A_l^{kj}}{(1-s\beta_1)^l}+\sum_{l=1}^{L_k}\frac{\tilde{A}_l^{kj}}{(1+s|\gamma_1|)^l}\right), \tag{1.19}$$

where the coefficients A_l^{kj} and $\tilde{A}_l^{kj}$ are given in (3.55) and are obtained by applying a partial fraction decomposition in (1.18) to arrive at (1.19). Substituting (1.19) in (1.13) and taking the inverse Laplace transform of $M_Y(-s)$, we obtain the PDF of Y as shown in (3.51).

References

[1] C. Zhang, S. Ariyavisitakul, and M. Tao, "LTE-Advanced and 4G wireless communications [guest editorial]," *IEEE Communications Magazine*, vol. 50, no. 2, pp. 102–103, February 2012.

[2] P. Bhat, S. Nagata, L. Campoy, I. Berberana, T. Derham, G. Liu, X. Shen, P. Zong, and J. Yang, "LTE-Advanced: An operator perspective," *IEEE Communications Magazine*, vol. 50, no. 2, pp. 104–114, February 2012.

[3] S. Chen, J. Zhao, and Y. Peng, "The development of TD-SCDMA 3G to TD-LTE-Advanced 4G from 1998 to 2013," *IEEE Wireless Communications*, vol. 21, no. 6, pp. 167–176, December 2014.

[4] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[5] Cisco, "Global mobile data traffic forecast update, 2013-2018," *White Paper*, 2014.

[6] Nokia Siemenes Networks, "2020: Beyond 4G radio evolution for the gigabit experience," *White Paper*, 2011.

[7] Ericsson, "More than 50 billion connected devices," *White Paper*, 2011.

[8] D. Gonzalez, A. Rusu, and M. Ismail, "Tackling 4G challenges with 'TACT' - Design and optimization of 4G radio receivers with a transceiver architecture comparison tool (TACT)," *IEEE Circuits and Devices Magazine*, vol. 22, no. 6, pp. 16–23, November 2006.

[9] P. Mak, U. Seng-Pan, and R. Martins, "Transceiver architecture selection: Review, state-of-the-art survey and case study," *IEEE Circuits and Systems Magazine*, vol. 7, no. 2, pp. 6–25, 2007.

[10] A. Baschirotto, R. Castello, F. Campi, G. Cesura, M. Toma, R. Guerrieri, A. Lodi, L. Lavagno, and P. Malcovati, "Baseband analog front-end and digital back-end for reconfigurable multi-standard terminals," *IEEE Circuits and Systems Magazine*, vol. 6, no. 1, pp. 8–28, First Quarter 2006.

[11] K. Sahota, "RF front-end requirements for 3G and beyond," in *Proc. IEEE Ultrasonics Symposium (IUS)*, October 2010, pp. 86–90.

[12] A. Abidi, "Direct-conversion radio transceivers for digital communications," *IEEE Journal of Solid-State Circuits*, vol. 30, no. 12, pp. 1399–1410, December 1995.

[13] B. Razavi, "Design considerations for direct-conversion receivers," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 44, no. 6, pp. 428–435, June 1997.

[14] S. Mirabbasi and K. Martin, "Classical and modern receiver architectures," *IEEE Communications Magazine*, vol. 38, no. 11, pp. 132–139, November 2000.

[15] J. Atallah and M. Ismail, "Future 4G front-ends enabling smooth vertical handovers," *IEEE Circuits and Devices Magazine*, vol. 22, no. 1, pp. 6–15, January 2006.

[16] J. L. Mehta, "Transceiver architectures for wireless ICs," *RF Design*, pp. 76–96, February 2001.

[17] L. Smaini, *RF Analog Impairments Modeling for Communication Systems Simulation: Application to OFDM-based Transceivers*. Wiley, 2012.

[18] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, February 2014.

[19] R. Enrico, *Phase noise and Frequency Stability in Oscillators*. Cambridge University Press, 2009.

[20] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. A. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for 5G mobile and wireless communications: the vision of the METIS project," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 26–35, 2014.

[21] R. Walden, "Analog-to-digital converter survey and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 4, pp. 539–550, April 1999.

[22] L. Bin, T. Rondeau, J. Reed, and C. Bostian, "Analog-to-digital converters," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 69–77, November 2005.

[23] B. Jonsson, "A survey of A/D-converter performance evolution," in *Proc. 17th IEEE International Conference on Electronics, Circuits, and Systems (ICECS)*, December 2010, pp. 766–769.

[24] T. Sundstrom, B. Murmann, and C. Svensson, "Power dissipation bounds for high-speed Nyquist analog-to-digital converters," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, no. 3, pp. 509–518, March 2009.

[25] T. Souders, D. R. Flach, C. Hagwood, and G. Yang, "The effects of timing jitter in sampling systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 39, no. 1, pp. 80–85, February 1990.

[26] Y. C. Jenq, "Digital spectra of nonuniformly sampled signals: Fundamentals and high-speed waveform digitizers," *IEEE Transactions on Instrumentation and Measurement*, vol. 37, no. 2, pp. 245–251, June 1988.

[27] ——, "Digital-to-analog (D/A) converters with nonuniformly sampled signals," *IEEE Transactions on Instrumentation and Measurement*, vol. 45, no. 1, pp. 56–59, February 1996.

[28] U. Seng-Pan, S. Sai-Weng, and R. Martins, "Exact spectra analysis of sampled signals with jitter-induced nonuniformly holding effects," *IEEE Transactions on Instrumentation and Measurement*, vol. 53, no. 4, pp. 1279–1288, 2004.

[29] U. Onunkwo, Y. Li, and A. Swami, "Effect of timing jitter on OFDM-based UWB systems," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 4, pp. 787–793, April 2006.

[30] B. Putra and G. Fettweis, "Clock jitter estimation and suppression in OFDM systems employing bandpass $\Sigma\Delta$ ADC," in *Proc. 10th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2009, pp. 623–627.

[31] ——, "The effect of clock jitter on the performance of bandpass $\Sigma\Delta$ ADCs," in *Proc. 3rd International Symposium on Communications, Control and Signal Processing (ISCCSP)*, March 2008, pp. 1334–1338.

[32] V. Syrjälä and M. Valkama, "Jitter mitigation in high-frequency bandpass-sampling OFDM radios," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, April 2009, pp. 1–6.

[33] ——, "Sampling jitter cancellation in direct-sampling radio," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, April 2010, pp. 1–6.

[34] M. Windisch and G. Fettweis, "On the performance of standard-independent I/Q imbalance compensation in OFDM direct-conversion receivers," in *Proc. 13th European Signal Processing Conference (EUSIPCO)*, 2005, pp. 4–8.

[35] A. Tarighat, R. Bagheri, and A. Sayed, "Compensation schemes and performance analysis of IQ imbalances in OFDM receivers," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 3257–3268, August 2005.

[36] L. Anttila, M. Valkama, and M. Renfors, "Blind compensation of frequency-selective I/Q imbalances in quadrature radio receivers: Circularity-based approach," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 3, April 2007, pp. 245–248.

[37] L. Anttila, "Digital front-end signal processing with widely-linear signal models in radio devices," PhD, Tampere University of Technology, Department of Communications Engineering, 2011.

[38] M. Y. Cheong, "Development of digital predistorters for broadband power amplifiers in OFDM systems using the simplical canonical piecewise linear function," PhD, Aalto University, School of Electrical Engineering, Dept. of Signal Processing and Acoustics, 2014.

[39] A. H. Gokceoglu, "Performance analysis and mitigation of nonlinear distortion, IQ imbalance and phase noise in modern radio communication disciplines," PhD, Tampere University of Technology, Department of Electronics and Communications Engineering, 2014.

[40] H. Ochiai and H. Imai, "On the distribution of the peak-to-average power ratio in OFDM signals," *IEEE Transactions on Communications*, vol. 49, no. 2, pp. 282–289, February 2001.

[41] S. Miller and R. O'Dea, "Peak power and bandwidth efficient linear modulation," *IEEE Transactions on Communications*, vol. 46, no. 12, pp. 1639–1648, December 1998.

[42] S. C. Cripps, *RF Power Amplifiers for Wireless Communications*, 2nd ed. Artech House.

[43] A. Tehrani, H. Cao, S. Afsardoost, T. Eriksson, M. Isaksson, and C. Fager, "A comparative analysis of the complexity/accuracy tradeoff in power amplifier behavioral models," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 6, pp. 1510–1520, June 2010.

[44] M. Isaksson, D. Wisell, and D. Ronnow, "A comparative analysis of behavioral models for RF power amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 1, pp. 348–359, January 2006.

[45] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Communications Magazine*, vol. 28, 1990.

[46] M. O. Pun, M. Morelli, and C. C. J. Kuo, *Multi-carrier techniques for broadband wireless communications*. Imperial College Press, 2007.

[47] 3rd Generation Partnership Project, "3GPP TS 36.104," European Telecommunications Standards Institute, Tech. Rep. version 13.1.0, 2015.

[48] ——, "3GPP TS 36.211," European Telecommunications Standards Institute, Tech. Rep. version 12.7.0, 2015.

[49] A. Sahin, I. Guvenc, and H. Arslan, "A survey on multicarrier communications: Prototype filters, lattice structures, and implementation aspects," *IEEE Communications Surveys Tutorials*, vol. 16, no. 3, pp. 1312–1338, 2014.

[50] S. Thompson, A. Ahmed, J. Proakis, J. Zeidler, and M. Geile, "Constant envelope OFDM," *IEEE Transactions on Communications*, vol. 56, no. 8, pp. 1300–1312, August 2008.

[51] J. Stott, "The effects of phase noise in COFDM," *EBU Technical Review*, Summer 1998.

[52] T. Schenk, *RF Imperfections in High-rate Wireless Systems*. Springer, 2008.

[53] S. Haykin and M. Moher, *Communication Systems*. John Wiley and Sons, 2009.

[54] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase noise in oscillators: A unifying theory and numerical methods for characterization," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 47, no. 5, pp. 655–674, May 2000.

[55] E. Rubiola, "The measurement of AM noise of oscillators," in *Proc. IEEE International Frequency Control Symposium and Exposition*, June 2006, pp. 750–758.

[56] P. Maffezzoni, F. Pepe, and A. Bonfanti, "A unified method for the analysis of phase and amplitude noise in electrical oscillators," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 9, pp. 3277–3284, September 2013.

[57] T. Lee and A. Hajimiri, "Oscillator phase noise: A tutorial," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 3, pp. 326–336, March 2000.

[58] A. Chorti and M. Brookes, "A spectral model for RF oscillators with power-law phase noise," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 53, no. 9, pp. 1989–1999, September 2006.

[59] M. Khanzadi, D. Kuylenstierna, A. Panahi, T. Eriksson, and H. Zirath, "Calculation of the performance of communication systems from measured oscillator phase noise," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, no. 5, pp. 1553–1565, May 2014.

[60] G. Klimovitch, "Near-carrier oscillator spectrum due to flicker and white noise," in *Proc. IEEE International Symposium on Circuits and Systems*, vol. 1, May 2000, pp. 703–706.

[61] S. Sancho, A. Suarez, J. Dominguez, and F. Ramirez, "Analysis of near-carrier phase-noise spectrum in free-running oscillators in the presence of white and colored noise sources," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 3, pp. 587–601, March 2010.

[62] B. Razavi, "A study of phase noise in CMOS oscillators," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 3, pp. 331–343, March 1996.

[63] A. Demir, "Phase noise and timing jitter in oscillators with colored-noise sources," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 49, no. 12, pp. 1782–1791, December 2002.

[64] A. Mehrotra, "Noise analysis of phase-locked loops," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 49, no. 9, pp. 1309–1316, September 2002.

[65] C. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and Natural Sciences*. Springer-Verlag, 1994.

[66] F. M. Gardner, "Charge-pump phase-lock loops," *IEEE Transactions on Communications*, vol. 28, no. 11, pp. 1849–1858, November 1980.

[67] P. Hanumolu, M. Brownlee, K. Mayaram, and U. Moon, "Analysis of charge-pump phase-locked loops," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 51, no. 9, pp. 1665–1674, September 2004.

[68] P. Maffezzoni and S. Levantino, "Analysis of VCO phase noise in charge-pump phase-locked loops," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 59, no. 10, pp. 2165–2175, October 2012.

[69] J. G. Proakis and M. Salehi, *Digital communications*, 5th ed. McGraw-Hill, 2008.

[70] H. Sari, G. Karam, and I. Jeanclaude, "Channel equalization and carrier synchronization in OFDM systems," in *Proc. 6th International Symposium on Workshop on Digital Communications*, September 1993, pp. 191–202.

[71] H. Nikookar and R. Prasad, "Performance evaluation of multi-carrier transmission over measured indoor radio propagation channels," in *Proc. 6th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Wireless: Merging onto the Information Superhighway*, vol. 1, September 1995, pp. 61–65.

[72] A. Armada and M. Calvo, "Phase noise and sub-carrier spacing effects on the performance of an OFDM communication system," *IEEE Communications Letters*, vol. 2, no. 1, pp. 11–13, January 1998.

[73] A. Armada, "Understanding the effects of phase noise in orthogonal frequency division multiplexing (OFDM)," *IEEE Transactions on Broadcasting*, vol. 47, no. 2, pp. 153–159, June 2001.

[74] T. Pollet, M. Van Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Transactions on Communications*, vol. 43, no. 234, pp. 191–193, February/April 1995.

[75] H. Nishookar and R. Prasad, "On the sensitivity of multicarrier transmission over multipath channels to phase noise and frequency offset," in *Proc. 17th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, vol. 1, October 1996, pp. 68–72.

[76] H. G. Ryu, Y. S. Li, and J. S. Park, "Nonlinear analysis of the phase noise in the OFDM communication system," *IEEE Transactions on Consumer Electronics*, vol. 50, no. 1, pp. 54–63, February 2004.

[77] W. Songping and Y. Bar-Ness, "OFDM systems in the presence of phase noise: Consequences and solutions," *IEEE Transactions on Communications*, vol. 52, no. 11, pp. 1988–1996, November 2004.

[78] L. Piazzo and P. Mandarini, "Analysis of phase noise effects in OFDM modems," *IEEE Transactions on Communications*, vol. 50, no. 10, pp. 1696–1705, October 2002.

[79] L. Tomba, "On the effect of Wiener phase noise in OFDM systems," *IEEE Transactions on Communications*, vol. 46, no. 5, pp. 580–583, May 1998.

[80] M. El-Tanany, Y. Wu, and L. Hazy, "Analytical modeling and simulation of phase noise interference in OFDM-based digital television terrestrial broadcasting systems," *IEEE Transactions on Broadcasting*, vol. 47, no. 1, pp. 20–31, March 2001.

[81] K. Sathananthan and C. Tellambura, "Performance analysis of an OFDM system with carrier frequency offset and phase noise," in *Proc. 54th IEEE Vehicular Technology Conference (VTC)*, vol. 4, Fall 2001, pp. 2329–2332.

[82] ——, "Probability of error calculation of OFDM systems with frequency offset," *IEEE Transactions on Communications*, vol. 49, no. 11, pp. 1884–1888, November 2001.

[83] T. Schenk, R. van der Hofstad, E. Fledderus, and P. Smulders, "Distribution of the ICI term in phase noise impaired OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1488–1500, April 2007.

[84] K. Hamdi, "Exact SINR analysis of wireless OFDM in the presence of carrier frequency offset," *IEEE Transactions on Wireless Communications*, vol. 9, no. 3, pp. 975–979, March 2010.

[85] E. Costa and S. Pupolin, "M-QAM-OFDM system performance in the presence of a nonlinear amplifier and phase noise," *IEEE Transactions on Communications*, vol. 50, no. 3, pp. 462–472, March 2002.

[86] J. Montojo and L. Milstein, "Effects of imperfections on the performance of OFDM systems," *IEEE Transactions on Communications*, vol. 57, no. 7, pp. 2060–2070, July 2009.

[87] G. Fettweis, M. Löhning, D. Petrovic, M. Windisch, P. Zillmann, and W. Rave, "Dirty RF: A new paradigm," *International Journal of Wireless Information Networks*, vol. 14, no. 2, pp. 133–148, June 2007.

[88] R. Stuhlberger, R. Krueger, B. Adler, J. Kissing, L. Maurer, G. Hueber, and A. Springer, "LTE-downlink performance in the presence of RF-impairments," in *Proc. European Conference on Wireless Technologies*, October 2007, pp. 189–192.

[89] M. Jalloh and P. Das, "Performance analysis of STBC-OFDM transmit diversity with phase noise and imperfect channel estimation," in *Proc. IEEE Military Communications Conference*, November 2008, pp. 1–5.

[90] S. Bittner, M. Krondorf, and G. Fettweis, "Numerical performance evaluation of OFDM systems affected by transmitter nonlinearities, phase noise and channel estimation errors," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, November/December 2008.

[91] V. Syrjälä, "Analysis and mitigation of oscillator impairments in modern receiver architectures," PhD, Tampere University of Technology. Dept. of Communications Engineering, 2012.

[92] A. Gokceoglu, Y. Zou, M. Valkama, P. Sofotasios, P. Mathecken, and D. Cabric, "Mutual information analysis of OFDM radio link under phase noise, IQ imbalance and frequency-selective fading channel," *IEEE Transactions on Wireless Communications*, vol. 12, no. 6, pp. 3048–3059, June 2013.

[93] "Understanding phase noise needs and choices in signal generation," Keysight Technologies, Tech. Rep., 2014.

[94] "LTE system specifications and their impact on RF and base band circuits," Rohde and Schwarz, Tech. Rep., 2013.

[95] P. Moschopoulos, "The distribution of the sum of independent gamma random variables," *Annals of the Institute of Statistical Mathematics*, vol. 37, no. 1, pp. 541–544, December 1985.

[96] M. S. Alouini, A. Abdi, and M. Kaveh, "Sum of gamma variates and performance of wireless communication systems over Nakagami-fading channels," *IEEE Transactions on Vehicular Technology*, vol. 50, no. 6, pp. 1471–1480, November 2001.

[97] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products*. Academic Press, 2007.

[98] S. Krone and G. Fettweis, "Capacity analysis for OFDM systems with transceiver I/Q imbalance," in *Proc. IEEE Global Telecommunications Conference*, November 2008, pp. 1–6.

[99] R. Krishnan, G. Colavolpe, A. Graell i Amat, and T. Eriksson, "Algorithms for joint phase estimation and decoding for MIMO systems in the presence of phase noise and quasi-static fading channels," *IEEE Transactions on Signal Processing*, vol. 63, no. 13, pp. 3360–3375, July 2015.

[100] R. Krishnan, M. Khanzadi, T. Eriksson, and T. Svensson, "Soft metrics and their performance analysis for optimal data detection in the presence of strong oscillator phase noise," *IEEE Transactions on Communications*, vol. 61, no. 6, pp. 2385–2395, June 2013.

[101] H. Mehrpouyan, A. Nasir, S. Blostein, T. Eriksson, G. Karagiannidis, and T. Svensson, "Joint estimation of channel and oscillator phase noise in MIMO systems," *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4790–4807, September 2012.

[102] G. Colavolpe, A. Barbieri, and G. Caire, "Algorithms for iterative decoding in the presence of strong phase noise," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 9, pp. 1748–1757, September 2005.

[103] M. Nissila and S. Pasupathy, "Adaptive iterative detectors for phase-uncertain channels via variational bounding," *IEEE Transactions on Communications*, vol. 57, no. 3, pp. 716–725, March 2009.

[104] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge University Press, 2013.

[105] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*. Plenum Press, 1997.

[106] J. Bingham, "Method and apparatus for correcting for clock and carrier frequency offset, and phase jitter in multicarrier modems," Patent US 5206 886, April 13, 1993.

[107] R. A. Casas, S. L. Biracree, and A. E. Youtz, "Time domain phase noise correction for OFDM signals," *IEEE Transactions on Broadcasting*, vol. 48, no. 3, pp. s230–236, September 2002.

[108] S. Kay, *Fundamentals of statistical signal processing: Estimation theory*. Prentice Hall, 1993.

[109] S. Wu, P. Liu, and Y. Bar-Ness, "Phase noise estimation and mitigation for OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 5, no. 12, pp. 3616–3625, December 2006.

[110] D. Petrovic, W. Rave, and G. Fettweis, "Effects of phase noise on OFDM systems with and without PLL: Characterization and compensation," *IEEE Transactions on Communications*, vol. 55, no. 8, pp. 1607–1616, August 2007.

[111] K. Nikitopoulos and A. Polydoros, "Phase-impairment effects and compensation algorithms for OFDM systems," *IEEE Transactions on Communications*, vol. 53, no. 4, pp. 698–707, April 2005.

[112] K. Kim, Q. Zou, H. J. Choi, and A. Sayed, "An efficient carrier phase synchronization technique for high-order M-QAM-OFDM," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3789–3794, August 2008.

[113] F. Munier, T. Eriksson, and A. Svensson, "An ICI reduction scheme for OFDM system with phase noise over fading channels," *IEEE Transactions on Communications*, vol. 56, no. 7, pp. 1119–1126, July 2008.

[114] V. Syrjälä, M. Valkama, N. Tchamov, and J. Rinne, "Phase noise modelling and mitigation techniques in OFDM communications systems," in *Proc. Wireless Telecommunications Symposium*, April 2009, pp. 1–7.

[115] V. Syrjälä and M. Valkama, "Analysis and mitigation of phase noise and sampling jitter in OFDM radio receivers," *International Journal of Microwave and Wireless Technologies*, vol. 2, no. 2, pp. 193–202, April 2010.

[116] ——, "Iterative receiver signal processing for joint mitigation of transmitter and receiver phase noise in OFDM-based cognitive radio link," in *Proc. International Conference on Cognitive Radio Oriented Wireless Networks*, June 2012.

[117] ——, "Receiver DSP for OFDM systems impaired by transmitter and receiver phase noise," in *Proc. IEEE International Conference on Communications (ICC)*, June 2011, pp. 1–6.

[118] D. D. Lin, R. Pacheco, T. J. Lim, and D. Hatzinakos, "Joint estimation of channel response, frequency offset, and phase noise in OFDM," *IEEE Transactions on Signal Processing*, vol. 54, no. 9, pp. 3542–3554, 2006.

[119] P. Rabiei, W. Namgoong, and N. Al-Dhahir, "A non-iterative technique for phase noise ICI mitigation in packet-based OFDM systems," *IEEE Transactions on Signal Processing*, vol. 58, no. 11, pp. 5945–5950, November 2010.

[120] R. Carvajal, J. Aguero, B. Godoy, and G. Goodwin, "EM-based maximum-likelihood channel estimation in multicarrier systems with phase distortion," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 1, pp. 152–160, January 2013.

[121] F. Septier, Y. Delignon, A. Menhaj-Rivenq, and C. Garnier, "Monte carlo methods for channel, phase noise, and frequency offset estimation with unknown noise variances in OFDM systems," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3613–3626, August 2008.

[122] Z. Wang and G. Giannakis, "Wireless multicarrier communications," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 29–48, May 2000.

[123] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. John Wiley and Sons, 2008.

[124] O. H. Salim, A. A. Nasir, H. Mehrpouyan, W. Xiang, S. Durrani, and R. A. Kennedy, "Channel, phase noise, and frequency offset in OFDM systems: Joint estimation, data detection, and hybrid Cramer-Rao lower bound," *IEEE Transactions on Communications*, vol. 62, no. 9, pp. 3311–3325, Sept 2014.

[125] D. D. Lin and T. J. Lim, "The variational inference approach to joint data detection and phase noise estimation in OFDM," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 1862–1874, 2007.

[126] M. J. Beal, "Variational algorithms for approximate Bayesian inference," PhD, Gatsby Computational Neuroscience Unit, University College London, 2003.

[127] D. MacKay, *Information theory, Inference and Learning Algorithms*. Cambridge University Press, 2003.

[128] P. Mathecken, T. Riihonen, S. Werner, and R. Wichman, "Constrained phase noise estimation in OFDM using scattered pilots without decision feedback," *Journal manuscript submitted for review*, 2016.

[129] E. Yonina Chana, "Quantum signal processing," PhD, Massachusetts Institute of Technology. Dept. of Electrical Engineering and Computer Science, 2002.

[130] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[131] Y. Nesterov and A. Nemirovski, *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, 1994.

[132] I. Polik and T. Terlaky, "A survey of the S-lemma," *SIAM Review*, vol. 49, pp. 371–418, 2007.

[133] V. A. Yakubovich, "S-procedure in nonlinear control theory," *Vestnik Leningrad. Univ. (English translation)*, vol. 4, pp. 73–93, 1977.

[134] U. T. Jonsson, *A Lecture on the S-procedure*. Division of Optimization and System Theory, Royal Institute of Technology (KTH), 2006.

[135] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, no. 1, pp. 49–95, March 1996.

[136] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization: Analysis, algorithms, and engineering applications*. SIAM, 2001.

[137] L. N. Trefethen and D. Bau, *Numerical Linear Algebra*. SIAM, 1997.

[138] A. Puglielli, A. Townley, G. LaCaille, V. Milovanovic, P. Lu, K. Trotskovsky, A. Whitcombe, N. Narevsky, G. Wright, T. Courtade, E. Alon, B. Nikolic, and A. M. Niknejad, "Design of energy- and cost-efficient massive MIMO arrays," *Proceedings of the IEEE*, vol. 104, no. 3, March 2016.

[139] T. Höhne and V. Ranki, "Phase noise in beamforming," *IEEE Transactions on Wireless Communications*, vol. 9, no. 12, pp. 3682–3689, December 2010.

[140] R. Krishnan, M. R. Khanzadi, N. Krishnan, Y. Wu, A. G. i Amat, T. Eriksson, and R. Schober, "Linear massive MIMO precoders in the presence of phase noise - A large-scale analysis," *IEEE Transactions on Vehicular Technology*, vol. PP, no. 99, pp. 1–1, 2015.

[141] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design," *IEEE Transactions on Wireless Communications*, vol. 14, no. 8, pp. 4353–4368, 2015.

[142] P. Banelli, S. Buzzi, G. Colavolpe, A. Modenini, F. Rusek, and A. Ugolini, "Modulation formats and waveforms for 5G networks: Who will be the heir of OFDM?: An overview of alternative modulation schemes for improved spectral efficiency," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 80–93, Nov 2014.

[143] M. Morelli, C. C. J. Kuo, and M. O. Pun, "Synchronization techniques for orthogonal frequency division multiple access (OFDMA): A tutorial review," *Proceedings of the IEEE*, vol. 95, no. 7, pp. 1394–1427, 2007.

[144] B. Farhang-Boroujeny, "OFDM versus filter bank multicarrier," *IEEE Signal Processing Magazine*, vol. 28, no. 3, pp. 92–112, 2011.

[145] K. S. Miller, *Multidimensional Gaussian Distributions*. John Wiley and Sons, 1964.

Aalto-DD 105/2016

ISBN 978-952-60-6838-1 (printed)
ISBN 978-952-60-6839-8 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)

Aalto University
School of Electrical Engineering
Department of Signal Processing and Acoustics
www.aalto.fi

BUSINESS +
ECONOMY

ART +
DESIGN +
ARCHITECTURE

SCIENCE +
TECHNOLOGY

CROSSOVER

DOCTORAL
DISSERTATIONS